2008

Concept relation extraction using natural language processing: the CRISP technique

Mohammed Abdelrahman Al Qady
Iowa State University

Concept relation extraction using natural language processing – the CRISP technique

by

Mohammed Abdelrahman Al Qady

A thesis submitted to the graduate faculty

in partial fulfillment of the requirements for the degree of

MASTER OF SCIENCE

Major: Civil Engineering (Construction Engineering and Management)

Program of Study Committee:
Amr Kandil, Major Professor
Edward Jaselskis
Nick Pendar

Iowa State University

Ames, Iowa

2008

TABLE OF CONTENTS

List of Tables

List of Figures

Acknowledgements

The research team would like to acknowledge the University of Utah's School of Computing for developing Sundance. Special thanks to Ellen Riloff and Siddharth Patwardhan from the University of Utah for giving us a copy of Sundance and for their valuable assistance.

Abstract

Knowledge Management (KM) has become the focus of a lot of scientific research during the second half of the twentieth century as researchers discovered the importance of the knowledge resource to business organizations. Recent research developed ontology-based knowledge management systems (KMS) to provide a standardized reference for knowledge consistency. However, use of ontologies has been impeded by the difficulties encountered in building ontologies, especially difficulties in the knowledge acquisition stage. It is hypothesized that NLP tools can be usefully implemented to assist in the knowledge acquisition stage for ontology building in specific, and to develop effective KMS's in general. The proposed system, CRISP, utilizes a shallow parser for extracting concept relations from construction contract documents to assist in the development of an ontology-based KMS. When compared with human evaluators, CRISP achieved almost 80% of the average kappa score attained by the evaluators, and approximately 90% of their F-measure score.

Chapter 1. Introduction

Overview

Knowledge Management (KM) has become the focus of a lot of scientific research during the second half of the twentieth century as researchers discovered the importance of the knowledge resource to business organizations. This importance is demonstrated throughout the various organizational levels by the emergence and use of terms such as “the knowledge society” and “the knowledge worker” (Drucker 1993), and the development of concepts such as “the learning organization” (Senge 1990). The way an organization manages knowledge involves how knowledge is created or extracted in the organization, preserved and communicated for effective utilization (Chinowsky and Molenaar 2005, Turk 2007, Walters et al. 2007).

In order to survive and succeed in gaining an advantage in the knowledge-intensive and highly competitive construction industry, effective use of large amounts of knowledge from various knowledge sources is essential. Realizing this fact, many construction firms started adopting and implementing some form of Knowledge Management System (KMS). However, many researchers agree that the current applications of KMS have not been able to achieve their full potential due to various reasons such as:

- The bureaucratic nature of construction organizations.
- Ineffective design of the KMS.
- Poor communication within the organization.

In general, research in KM can take two perspectives; the functionalist approach and the interpretive approach (Venters 2001). The functionalist perspective objectifies knowledge to facilitate its use. The majority of research in KM adopts this approach. The interpretive perspective views knowledge as the product of social processes and accordingly aims at promoting social interaction within the organization. Some researchers attempted to develop a KMS that combines both the functionalist and the interpretive approaches (Wetherill et al. 2007).

Recently, researchers have attempted to utilize the semantic relations in knowledge sources to improve KM techniques (Rezgui 2001). An ontology was used as a semantic reference to provide knowledge consistency. An ontology is a formal, explicit specification of a shared conceptualization (Borst 1997, Gruber 1993, Studer et al. 1998). Ontologies have been utilized in various applications such as e-commerce and the semantic web, and in diverse fields including construction engineering and law. In addition to standardizing the concepts used in a certain domain, ontologies can be used to facilitate natural language communications between humans and computers (El-Diraby and Kashif 2005).

Despite these advantages, the use of ontologies is limited for several reasons (Shamsfard and Barforoush 2004):

- Unavailability of standards for linking existing ontologies to enable sharing and reuse.
- Consensual agreement over the concepts of a certain phenomena (a key requirement for ontology development) is hard to achieve.
- The difficult and time-consuming stage of knowledge acquisition required for developing ontologies.

Research in the legal field attempted to use Natural Language Processing (NLP) tools to assist in developing legal ontologies from textual documents. In construction contract administration, the majority of contract documents is textual documents expressed in natural language. Contract documents express the agreement between the contracting parties on the rights and duties of each party. Accordingly, with regards to the parties of the contract, contract documents express consensual knowledge. An ontology that represents this consensual knowledge can be the center of an effective KMS used for project management functions such as contract administration and correspondence management.

Objective

It is hypothesized that NLP tools can be usefully implemented to assist in the knowledge acquisition stage for ontology building in specific, and to develop effective KMS's in general. This research proposes the use of the system CRISP (Concept Relation Identification using Shallow Parsing) for the automatic/semi-automatic extraction of concepts and concept relations from the text of contract documents. Although the system was originally developed for application in KMS's of construction firms to facilitate contract administration processes, CRISP can be used on any textual documents and is therefore not limited to a specific domain.

Thesis Organization

Chapter 2 gives an overview of knowledge and KM, a description of the use of KM in the construction field and recent research trends in KM. Chapter 3 describes the methodology used for developing CRISP and Chapter 4 presents the results of the evaluation of CRISP. Based on this evaluation, Chapter 5 offers several improvements for the development of CRISP in future work and lists applications that can benefit from the use of a system like CRISP.

Chapter 2. Literature Review

Overview

In today's knowledge society, knowledge is considered the key asset among the various assets of an organization. Most construction firms implement KM, however researchers believe that such implementations are immature and do not realize the full benefits that can be attained by the effective utilization of knowledge. Recent research developed ontology-based KMS to provide a standardized reference for knowledge consistency. However, the extensive use of ontologies has been impeded by the difficulties encountered in building ontologies, especially those difficulties related to the knowledge acquisition stage. Natural Language Processing tools have been previously used to assist in building legal ontologies. It is proposed to use NLP tools to assist in knowledge acquisition from construction contract documents in order to develop an effective ontology-based KMS for contract administration.

Knowledge and Organizations

In philosophy, knowledge is defined as justified true belief. The two traditional epistemological approaches to knowledge are rationalism which argues that knowledge is attained by a mental process through deductive reasoning, and empiricism which contends that knowledge is attained inductively through sensory experiences. The debate over the nature of knowledge has been ongoing for a long period of time dating back to ancient Greece, during which various philosophers have either championed one approach over the other or, more recently, attempted to combine both approaches (Nonaka and Takeuchi 1995).

Nonaka and Takeuchi (1995) identify two types of knowledge:

- Explicit knowledge: knowledge that has been explicitly explained, codified and expressed in formal language and can therefore by easily shared with others and effectively applied.
- Tacit knowledge: knowledge that is related to the intuition, perspective and experiences of an individual making its expression, representation and communication to others very difficult.

Examples of explicit knowledge in construction projects include the contract agreement, the specifications, reports, drawings, change orders; while tacit knowledge includes process records, problems faced, problems solved, expert suggestions, know-how, innovation, and experience notes (Lin et al. 2006).

The importance of knowledge to business organizations was recognized by researchers during the second half of the twentieth century, at which time the terms “knowledge society” and “knowledge worker” were coined by Peter Drucker. Drucker (1993) explains that the transformation into a knowledge society was the result of knowledge becoming the key business resource, not just another resource alongside the traditional resources of labor, capital and land. Various researchers share a similar view regarding the importance of KM: “Knowledge is the true asset of a marketing-oriented organization, and its integration across departments and disciplines should be emphasized” (Carneiro, 2001). The priority given to the knowledge asset over other assets in today’s business world is evidenced by the world-wide competition to gain and control knowledge and its means of communication (Toffler 1990). In the knowledge society, Drucker states that knowledge workers– workers that can successfully and efficiently utilize knowledge– are the key assets of an organization. In order to survive, organizations are required to develop and implement a continuous strategy of transformation in which outdated practices are discarded and new knowledge is created through innovation resulting from increased productivity of knowledge workers. This transformation is the essence of the learning organization model proposed by Senge (1990). A learning organization is characterized by implementing two kinds of learning:

- Adaptive learning: passive learning by which an organization adapts its processes in order to face new experiences and a changing environment.
- Generative learning: active learning by which an organization anticipates possible changes in the environment and creates new knowledge to face such changes.

Chinowsky and Molenaar (2005) recognize an organization's constant need to obtain knowledge as the primary driving force for achieving a proactive learning culture. The researchers developed a matrix that illustrates the general steps required in order to make the transformation into a learning organization. The entities involved in the learning process are listed across the top of the Learning Organization Matrix: the individual, the community of practice (a group of individual responsible for performing similar technical tasks), and the organization as a whole. The characteristics needed of the learning entities in order to implement the learning process are listed vertically: leadership, processes and infrastructure, communication, education, and finally culture. The researchers demonstrate how the matrix can be used as a maturity model to gauge the level of transformation of an organization into a learning organization. Case studies for ten organizations (both inside and outside the construction industry) were prepared and initial maturity models were developed for each case study to determine the level of implementation of learning organization.

Knowledge Management in Construction

In their book, The Knowledge-Creating Company, Nonaka and Takeuchi (1995) argue that the success of Japanese companies over the past few decades and their resilience in facing major economic crises is largely attributable to their ability to create new knowledge, i.e. the process of transforming tacit knowledge into explicit knowledge which is necessary for driving the innovation process. This book is accredited for being a cause of the increased interest by researchers in the field of knowledge management (Chinowsky and Molenaar 2005). Various definitions for knowledge management have been proposed (Raub and Ruling 2001). The majority of researchers in this field agree that knowledge management has to do with how an organization extracts/creates knowledge, and preserves it in a way that facilitates its communication throughout the organization in order to be utilized effectively in achieving organizational goals (Chinowsky and Molenaar 2005, Turk 2007, Walters et al. 2007).

Specifically for construction firms, Chinowsky and Carrillo (2007) argue that KM programs in engineering-procurement-construction (EPC) organizations must not be limited to the tasks of collection and distribution of knowledge, but must function as the foundation upon which the organizations develop into learning organizations, characterized by the objective of continually pursuing knowledge to enhance operations and acquire a competitive advantage. The objective of their research was to find the suitable connection necessary to allow this evolution by bridging the gap between KM and learning organization. Implementation of KM in an organization is modeled by the STEPS model consisting of 5 Stages: start-up stage, take-off stage, expansion stage, progressive stage and sustainability stage. By analyzing the results of case studies of 4 EPC organizations that actively pursue KM, the researchers concluded that a learning initiative can start at the expansion stage of the STEPS model, because it is at this stage that the KM initiative transforms from a project-based initiative to the organization-based initiative required to pursue a learning culture. After the expansion stage, development continues along two lines; towards sustainability in KM and maturity in learning organization. The maturity level of the learning organization can be categorized into 5 levels: establishing leadership, leadership transformation, integration of learning at both the community and individual levels, learning championing by leadership, and finally maturity. Although organizations may follow paths different from the roadmap suggested by the researchers and achieve the same end results, the analysis demonstrated that some of the alternative paths may delay or limit the ability to reach a learning culture. Such alternatives include attempting to reach a fully sustainable KM process before embarking on the learning initiative or trying to initiate learning at the early stage of KM (start-up and take-off).

Construction projects are described as complex and diverse projects that produce non-standard components (Clough et al. 2000). Yet there exists amongst this complexity and diversity vast amounts of knowledge. The construction industry is a perfect example of a knowledge-intensive industry that has offered and still has the ability to offer many contributions to research in the field of KM (Rezgui 2001). The dynamic business environment of the industry requires construction firms to effectively implement knowledge management systems in order to acquire a competitive edge in the market (Jung et al. 2006). How an organization manages knowledge has a significant impact on the organization's learning capabilities (Wetherill et al. 2007). These factors have led many construction firms to invest resources in developing effective KMS's. One study in the United Kingdom estimates that 40% of construction and design firms implement some type of KM strategy (Carrillo et al. 2004). A survey conducted with experts and engineers of the industry highlights the importance of KM and its benefits to construction projects (Lin et al. 2006).

Turk (2007) presents the results obtained from a survey of the European construction informatics research community aiming at identifying research topics in the field of construction informatics that have been addressed in the past, in addition to forecasting the research topics that will be focused on in the future. A map of research topics in construction informatics was developed and a set of research themes were identified (including, inter alia, concurrent engineering infrastructures, software interoperability and integration, KM, knowledge intensive applications). Opinions of the participants in the survey were polled in order to identify the current status of the research themes, their expected future trends, possibility of the migration of the researchers of a specific theme to other themes, and barriers encountered by researchers in the research themes. The results show an expected increase in KM research with possible migration to closely related themes such as information retrieval. Similarly, research in the theme of knowledge intensive applications (such as data mining, expert systems, and other applications employing artificial intelligence techniques) is expected to exhibit a positive trend in the future, with the possibility of migration of researchers in this field to KM and software interoperability/integration. The researcher states that these results, compounded by the current focus on software interoperability/integration (which according to the survey is also expected to continue), highlight the construction informatics community's belief that structural interoperability must exist not only on the syntactic level, but also on the semantic level, and thus building information models can be utilized for intelligent decision-making.

In another research, a case study of the opinions of industry experts was used to identify the trends and the future visions of the electrical contracting business (Walters et al. 2007). The results of telephone interviews with nine members of the National Electrical Contractors Association (NECA) were collected and classified according to the Process Classification Framework developed by the American Productivity and Quality Center (APQC) in order to facilitate cross-industry comparisons and consequently identify best practices in the industry that lead to best performance. Based on the outcome of the case study, the researchers conclude that a KM process is one of the important best practices that must be implemented by electrical contracting firms in order to stay competitive in the market and secure future success.

These studies demonstrate the importance of KM in the opinions of practitioners in the construction industry. Despite this importance and despite the wealth of knowledge generated in construction projects, the application of KM in construction is described as “immature and underutilized” (Asprey 2004, Laudon and Laudon 1998, McGee and Prusak 1993, Rezgui 2001). At the project level, many researchers agree that current KM practices fail to effectively utilize project-generated knowledge (McGee and Prusak 1993, Asprey 2004, Sor 2004). Reasons for this failure and proposed solutions are documented in various research.

Wetherill et al. (2007) comment on the authoritarian and bureaucratic nature of many construction organizations. In such organizations, knowledge creation, instead of being the job of all members of the organization, is the responsibility of a specific management group which consequently impedes KM activities.

Jung et al. (2006) attribute the failure of KMS’s to satisfy project needs to the design of the KMS. Many current KMS’s try to capture all related knowledge, regardless of its frequency and use, which is effort and time consuming. To overcome this, the researchers propose a methodology for analyzing the different knowledge areas of an organization to determine the best areas for developing a KMS. This is achieved at the macro level by a strategic fit analysis that starts with the highest level strategy, the corporate strategy, and identifies the engineering strategy which is consequently mapped onto the business functions. By comparing the required capability of the various business functions with the actual capabilities of the organization, those functions requiring the largest improvement are identified. The methodology continues with a disseminative fit analysis at the macro level, where knowledge sources for the identified business functions are listed and evaluated for the specific organization in terms of frequency (a measure of occurrence, accumulation and reuse of the knowledge source), explicitness (tacit or explicit data) and origin (obtained through individual or organizational learning). The current ‘As Is’ results are compared with future anticipated ‘To Be’ results in order to determine the appropriate knowledge sources that should be the subject of the KMS required to achieve the organization’s strategies. The proposed methodology was implemented on a general contracting firm specializing in building construction to identify the KMS areas necessary for the design management function.

Another method proposed for improving KM within an organization is to improve communication. The competitive and chaotic nature of the industry results in the breaking down of communication lines within an organization and between organizations (Walters et al. 2007). To counter this negative effect, organizations looked to adjusting their organizational hierarchy by:

- Flattening the hierarchical structure: Flatter structures allow more communications across functional boundaries. For example, project managers have more access to top management according to this approach.
- Adopting a matrix organization: This approach promotes interaction between employees, thus reinforcing communication.

Construction organizations have different perspectives regarding the KM issue. Whereas some organizations invest in developing an infrastructure for communicating knowledge, other organizations focus on developing and supporting social networks through which knowledge is shared (Chinowsky and Molenaar

2005). According to Venters (2001), approaches to KM generally fall under two perspectives:

- The functionalist perspective: This approach regards knowledge as an asset, and aims at objectifying knowledge in order to facilitate its storage, dissemination, retrieval and reuse. This approach is highly technological, emphasizing the use of information technology (IT), intranets and database systems to facilitate the access of users to required information. Although KM is not specifically an IT issue, IT is necessary to provide the framework for accessing and communicating knowledge (Chinowsky and Carrillo 2007). An example for the functionalist perspective is Nonaka and Takeuchi's approach explained above which categorizes knowledge as tacit and explicit knowledge and attempts to transform tacit knowledge into explicit knowledge.
- The interpretive perspective: This approach regards knowledge as being socially constructed and therefore must be viewed in relation to the social structure and processes of the organization. It is therefore the job of KM to promote the social context in which knowledge is shared by supporting the activities that knowledge workers participate in. As such, KM is expected to promote the development and growth of communities of practice. Under this perspective, technology may be used to assist social activities rather than to process data and information.

A lot of research in KM in construction took an approach from a functionalist perspective, fueled by the industry's need for a technological 'silver bullet' to address the KM issue (Chinowsky and Molenaar 2005). For example, Lin et al.

(2006) propose a methodology for capturing tacit and explicit knowledge during the construction phase of a project. In the Map-Based Knowledge Management system (MBKM), the knowledge acquired for the various construction activities is stored to enable knowledge re-use thereby reducing the time and cost of solving recurring problems, saving time by minimizing the need to refer to previous projects and providing better solutions for problems encountered in the construction phase. In the knowledge determination phase, the level of detail of the required knowledge is determined. The researchers recommend conducting the analysis at the level of detail of a construction activity. The knowledge extraction phase is where tacit and explicit knowledge related to a specific activity is identified. A senior project engineer collects information and documents for the activity and provides comments on the collected data. The next phase is the knowledge attribute phase, where the domain knowledge (including digital video, photographs, experts' opinions on problems, the senior engineer's description of the learned experience, etc.) is collected to build the knowledge attribute of the specific activity. In the knowledge linking phase, similar activities are linked, similar knowledge is linked, and knowledge is linked to the activities based on their relationship. Here, a knowledge worker classifies the knowledge by mapping the knowledge attribute onto the project's activities' map. Finally, the knowledge map is validated by domain experts, knowledge workers, etc., in the knowledge validation phase. Three kinds of search functions are supported by the MBKM system; project category search, keyword search and expert category search. The researchers note that the MBKM system should be utilized in successive projects and should be continually updated as more and more projects are executed in order build a database that captures and shares the experience and know-how attained by the project members in each individual project. The feedback obtained from the case-study in which the MBKM system was implemented shows that knowledge acquisition and extraction is a manual and time-consuming process. Senior engineers complained that the process of editing and recording information is inconvenient and takes time. Also, the search capabilities of the MBKM system might be problematic. As the system is updated, the amount of knowledge (documents, problem solution descriptions, etc.) available for a certain activity might become overwhelmingly large, thereby impeding easy access to relevant information in the system, and consequently defeating the original purpose for which the system was developed, that of providing knowledge sharing and re-use.

Another example for research with a functionalist approach is the Knowledge Document Management (KDM) Portal developed by Lee et al. (2005) with the objective of providing easy access to a construction project's technical documents. KDM is a knowledge management system that collects and organizes technical data in one location to facilitate document search and sharing. KDM includes a browser similar to the Windows Explorer browser that browses documents over three levels: documents located on the user's PC, shared documents based on sharing rights, and documents on the organization's server. According to the access rights, a user can review and share documents at the various levels. Documents are classified by the users into folders and sub-folders. A mutual connectivity option that connects similar documents is also available for the users. KDM Portal also provides search abilities for retrieving documents according to a specific search term. The research term is highlighted throughout the retrieved documents. The KDM portal allows the user to view different types of files (image files, AutoCAD, doc., xls., ppt.) even if the associated software is not available.

Some researchers recognized the increased benefits from having a KMS that not only focuses on manipulating formal knowledge but also supports developing social networks that increase knowledge creation and dissemination. Accordingly, some research attempted to combine both perspectives by providing a social aspect to the KMS's functions through facilitating and encouraging user interaction. An example for this type of research is the C-Sand project by Wetherill et al. (2007). The objective of the C-Sand project is twofold: to provide a framework that operates within and across organizations for the extraction and sharing of knowledge related to sustainability in construction, and to foster the necessary social functions required to develop such a framework. Interviews with industry professionals were conducted to identify current practices and industry needs which were then modeled in UML. Based on the UML diagrams and the above mentioned objectives, the architecture of the C-Sand system was developed. The C-Sand model is composed of nodes (representations of knowledge) and links between the nodes defining the relations between the knowledge representations. The C-Sand system provides several important KM features such as:

- Creating search 'interests' by developing a profile of the interests a user has previously searched and rated. The interests themselves can then be retrieved by other users in addition to tracing the user that created the interest, thereby allowing the users to contact an experienced person in the specific knowledge domain.

- A push function that allows users to flag a knowledge resource to others users with a description of the reason for flagging. A pull function by which users can subscribe to certain knowledge resources to receive updates from the system about recent changes in the resource.
- A spider function that crawls through a user defined URL searching for links to other documents and comparing those documents to the user's search criteria in order to extract additional relevant documents.

For evaluation, the C-Sand system was implemented and tested by the project's industry collaborators (various firms specializing in architecture, engineering design, building construction, quantity surveying and construction management). The results were generally positive, the major comments being on the interface design. The researchers state that although the system was designed for KM in the field of sustainable construction, the same approach can be adopted in other construction domains.

It is mentioned that the C-Sand project does not address the issue of creating links between knowledge resources based on their semantic relationship. A similar comment can be made on the KDM portal mentioned above. KDM uses simple search techniques, retrieving the documents that contain the search term regardless of the semantics behind the search query. In addition, data in the KDM is stored as-is, without any formal representation of the information in the documents which limits the program's ability to search and retrieve the relevant information. The researchers in the C-Sand project also state that the results of the system's evaluation express the need for conceptualizing the concepts of the research domain into a sustainability ontology by producing semantic representations of the knowledge resources and identifying semantic relations between the resources.

Rezgui (2001) proposes a knowledge environment that attempts to objectify knowledge by using information retrieval techniques, while also maintaining the social aspect of KM through user profiling. In this research, the proposed knowledge environment is centered on an ontology that is used as a semantic reference to provide knowledge consistency. In order to facilitate information retrieval, each document is represented by a semantic vector through the following process:

- Stop-words are removed and content words are stemmed.
- Index terms are extracted and prioritized according to their frequency in the document.
- To maintain knowledge consistency, index terms are mapped onto the ontology, either directly by identifying the corresponding concepts in the ontology, or indirectly by utilizing a thesaurus to identify the closest concepts in the ontology.
- The concepts are expanded according to ontology relationships, namely: generalization/specialization; composition/aggregation; concept association.
- Based on the term frequency index and the type of mapping used, a weighing factor is applied to each ontology concept to produce the semantic vector of the document.

In a similar fashion, user information and interests are mapped onto the ontology to create semantic vectors of user profiles, thus enabling searching through user profiles in a manner similar to document searches. The search service of the portal provides an advanced search option in which ontological concepts related to the search query are provided for the user to enhance the search.

Ontologies

Ontologies have been used in various fields including e-commerce, information retrieval, the semantic web and KM (Gómez-Pérez et al. 2004). In the field of Knowledge Engineering, an ontology could be defined as "an explicit specification of a conceptualization" (Gruber 1993), where 'conceptualization' refers to modeling an abstract world phenomenon by identifying the concepts pertaining to that phenomenon, and 'explicit' implies that the concepts, the relations between the concepts and the constraints on the concepts are explicitly defined (Studer et al. 1998). Borst (1997) added the word formal to this definition to emphasize that an ontology should be machine-readable, and described the conceptualization as being shared to denote that an ontology expresses consensual knowledge, not the opinion of a specific person. The following ontology components are evident from this definition:

- Concepts: usually expressed as classes in a hierarchy having certain attributes or properties.
- Relation: associations between concepts.
- Axioms: fundamental statements that are assumed to be true and are used to constrain concepts and their relations.

Depending on the components of an ontology and their level of detail, ontologies can be classified into light-weight ontologies, or taxonomies, where concepts are classified into a hierarchy defined by one type of relation (usually a sub-class relation) and heavy-weight ontologies, where concepts have defined properties and their relations are constrained by defined axioms. According to the subject of the conceptualization, ontologies can be classified into a hierarchy that includes:

- Top-level ontologies: ontologies representing general concepts that root terms in other ontologies should be linked to.
- Domain ontologies: ontologies that model concepts, relationships between concepts, activities, and fundamental principles in a particular domain.
- Application ontologies: ontologies that extend domain and task ontologies to model the knowledge in a specific application.

The typical reusability-usability tradeoff applies; ontologies that model specific tasks demonstrate high usability but are not easily reusable in other fields (refer to Figure 1) (Gómez-Pérez et al. 2004).

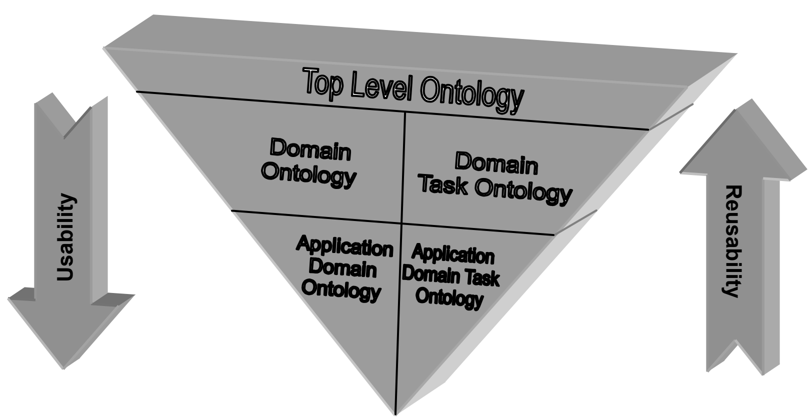

Figure 1: Sample ontology hierarchy

There are many advantages for modeling domains using ontologies. Ontologies support the development of common terms within a domain and the standardization of its concepts (Rezgui 2001). As an example from the construction industry, the huge amount of infrastructure rehabilitation projects requires that the numerous project participants share a common semantic basis for knowledge exchange. This will enhance collaboration between the various participants and will act as the backbone for any web services required for project communications. Expected advantages on the long run include the development of semantically compatible application software that will result in a positive impact on the efficiency of infrastructure projects' supply chains and a consequent reduction of the projects' indirect costs. El-Diraby and Kashif (2005) developed HiOnto, an ontology for highway construction and design processes. The architecture of HiOnto is composed of three levels connected by inheritance and mapping mechanisms. The upper most level, the domain level, consists of the e-COGNOS ontology, a process-oriented construction domain ontology developed during research in the area of Information Society Technologies (IST), part of the sixth program of the European Union's Framework Programme for Research and Technological Development. As an ontology, e-COGNOS has 3 main components: a taxonomy of concepts, relations between the concepts, and axioms bounding the concepts and their relations. Construction concepts in e-COGNOS are modeled as processes (PR), projects (PJ), products (PD), actors (AC), resources (RE) or technical topics (TT). e-COGNOS is described as process-centered because the various kinds of concepts are linked by a set of basic relations to some kind of engineering process concept;

for example, the relation ‘involves’ relates an AC to a certain PR, the relation ‘outputs’ relates a PD to a certain PR, the relation ‘utilizes’ relates an RE to a certain PR, the relation ‘constrained by’ relates a TT to a certain PR, etc. In addition to the basic relations, e-COGNOS contains axioms that define:

- The state of a concept at a certain time (temporal control axioms).
- The state of a PR concept as the aggregation of the states of its various related AC, RE and PD concepts.
- The state of a PJ concept as the aggregation of the states of its PR.
- The cause of change of a PR’s state and its sub-processes in the representation of the PR.

The second level in HiOnto, the application level, consists of a set of application ontologies for the same general concepts (project, process, product, actor, resource), however only as they apply to the specific field of highway construction. Each application ontology contains the 3 components of an ontology (taxonomy, relations and axioms) thereby producing 5x3 matrix. As in e-COGNOS, the different application ontologies are centered around the process ontology. Moreover, the application ontologies inherit relations and axioms from their parent concepts in the domain ontology. Technical topic ontologies are used to constrain the application ontologies. For example, the process of highway design in the process ontology, must comply with design codes, part of the technical topic ontologies. It is important to note that the process taxonomy in e-COGNOS was augmented with two additional processes, analysis process and testing process, to act as parent processes for the equivalent processes at the application level. Similarly, the actor’s taxonomy in e-COGNOS contained 2 major actors (organizations and personnel) and was modified by adding a third actor type, other actors, to take into account the broader list of actors available in highway construction.

The third level of HiOnto, the user level, represents actual instances of concepts from actual ongoing enterprises, in which taxonomies, relations and axioms inherited from the domain level into the application level are mapped onto actual enterprise models.

A taxonomy of the highway construction domain was developed based on the 'is-a' relation to allow interaction between the various application ontologies. A list of 4,000 terms related to highway construction was identified from technical texts (research papers, textbooks, construction handbooks) and interviews with experts. Using process-based competency questions (CQ), the terms were then categorized into the main e-COGNOS concepts and structured into a hierarchy. El-Diraby and Kashif (2005) also state another important advantage resulting from the use of ontologies. An ontology can be used to provide a representation of natural language that can be processed by a computer.

However, despite the advantages offered by ontologies, few research have attempted to utilize ontologies for KM in the construction industry. Shamsfard and Barforoush (2004) state several problems that limit the use of ontologies. Despite recent efforts, standards for the integration of existing ontologies to enable sharing and reuse are largely unavailable. In addition, consensual agreement on the concepts of a certain phenomenon is very hard to achieve, even if the phenomenon was limited to a very specific domain. Finally, the extraction of concepts and concept relations is a difficult and time-consuming process. This is demonstrated in HiOnto where knowledge acquisition required a lot of manual work. For example, the 4000 terms used to build the ontology's taxonomy were extracted from texts or identified from expert interviews. In addition, an object model of the domain based on the author's views had to be built in order to identify concept relations. For domains expressed in textual documents, the process of ontology building can be improved by utilizing natural language processing (NLP) techniques (Lame 2004).

Natural Language Processing

Speech and Language Processing is the field that aims at developing applications that utilize natural language processing, computational linguistics and speech recognition and synthesis in order to process human language as language, i.e. by means of the applications' use of knowledge of languages (Jurafsky and Martin 2000). The ultimate goal of research in natural language understanding is to produce computational models that resemble human beings in their linguistic abilities (reading, writing, hearing and speaking) (Allen 1995). In order to do so, knowledge from diverse scientific disciplines is required (linguistics, psycholinguistics, philosophy, computational linguistics). Aspects of NLP can be distinguished in Rezgui's (2001) ontology-centered knowledge environment discussed earlier, specifically in the process of developing the semantic vectors for the documents, such as: identification of document terms and their frequencies, stemming of terms, term frequency index calculation. The utilization of a lexical resource (thesaurus) is another NLP-related feature of this project, even if utilization was on a very limited scale (to indirectly map document terms onto their related ontological concept).

The origins of language processing can be traced back as early as the 1940's (Jurafsky and Martin 2000). The AI attribute of NLP was emphasized by the Turing Test (Turing 1950). This test (involving three participants in which a human interrogator tries to determine, by asking questions, which of the other two participants is a person and which is a machine) established that the ability to use language as humans do can be regarded as a measure of intelligence. The late nineties evidenced significant changes in the development of NLP techniques with the advent of probabilistic methods to refine the various algorithms (parsing, tagging, etc.), the emergence of the Web which emphasized the importance of language-based information extraction/retrieval and with the rapid advance in computer technology (Jurafsky and Martin 2000).

One of the earliest applications of NLP was the ELIZA program developed at MIT in the 1960's (Allen 1995). ELIZA plays the role of a therapist, asking questions based on the answers of the user, who plays the role of the patient. The program contains a database of keywords and a specification of output for each keyword. The program searches for a keyword in the user's answer and asks the following question based on the output specified for the keyword. ELIZA therefore does not actually understand the dialogue with the user, nor does it make any arguments, conclusions, or claims. This is acceptable in this particular dialogue between a therapist and a patient in which the therapist can pretend to not know anything about the real world (Jurafsky and Martin 2000).

Perhaps the most recognized uses for NLP techniques today are those related to commercial applications such as the spelling and grammar correcting capabilities of modern word processors (Church and Rau 1995). However, text-based NLP techniques have been utilized in numerous applications such as information extraction and retrieval, automatic text summarization and machine translation (Allen 1995). Such NLP-enabled applications have been used in various fields including financial analysis, computer software development and law.

In the financial field, the University of Durham developed a financial information extraction system that employs NLP techniques to extract and present to financial operators the relevant financial information available in source articles, thereby saving the time required to read the whole article (Costantino et al. 1997). This is achieved by syntactically and semantically comparing the article with predefined financial templates and mapping the information extracted from the article onto the most relevant template. The system also supports augmenting the predefined templates by adding user-defined templates. Paik et al. (2001) describe using <!metaMarker> an automatic metadata extraction system that utilizes machine learning and NLP techniques to process email communication between financial analysts and their clients in order to personalize the clients' profile according to the information extracted from the emails. Personalization of the profiles is not only done by using explicit information in the emails, but also by using implicit information that the system can infer as a result of being trained on thousands of emails.

The syntax of programming languages can be the cause of problems and difficulties for many programmers, whether professionals or beginners. NaturalJava is a prototype tool that provides programmers with a user interface for describing the required program using natural English language, and the program automatically produces the required Java source code (Price et al. 2000). The popularity of object-oriented modeling in software development resulted in many attempts to use NLP techniques to assist in the object-oriented analysis stage. Harman and Gaizauskas (2003) presented CM-Builder, a NL-based computer aided software engineering tool that uses NLP techniques to analyze software requirement texts written in English and develop an initial model for the object classes mentioned in the text and the relationships among them, expressed in UML (Unified Modeling Language). Similarly, Mala and Uma (2006) propose a system called Requirements Elicitor for the automatic construction of an object oriented design model (expressed in UML diagram) from requirement specifications (expressed in natural language).

In the legal domain, various attempts at using NLP techniques for information retrieval are documented. Hachey and Grover (2006) used NLP to perform automatic summarization of the judgments of the United Kingdom's House of Lords. The objective of their research was to provide applications to enhance legal information management by means of quick access to judgments' summaries and dynamic, customizable information retrieval.

Using NLP for Building Ontologies

Research in the legal field offers valuable insight on how NLP can be effectively applied to facilitate the process of developing ontologies. Lame (2004) describes how NLP techniques can be used to perform the conceptualization stage (identification of concepts and concept relations) in the process of building an ontology dedicated for information retrieval from a corpus of legal texts. The 57 codes of the French legal system are used as the corpus for this project. Since the function of law is to regulate the actions of people, law is considered to conceptualize the world. In addition, it is assumed that the legislator during the process of creating legal codes (codification) conceptualizes the legal domain into legal rules, with legal terms labeling legal concepts. Accordingly, ‘legal terms’ are not specifically terms related to the field of law, but include all worldly objects that must be captured by law and conceptualized by the legislator in codes in order to regulate human activities. A shallow syntactic parser, Syntex, was used to extract over 500,000 terms. In order to separate legal terms from non-legal terms, an initial filter was applied in which adjectives and adverbs were removed (on the assumption that concepts are expressed as nouns or noun phrases) and terms containing non-alphabetical characters were excluded, leaving a total count of approximately 300,000 terms. Statistical analysis was performed in an attempt to determine statistical indices’ thresholds that separate between legal and non-legal terms. The following statistical indices were used: term frequency (tf), inverse document frequency (idf), term frequency inverse document frequency (tf.idf) and entropy. No definite demarcation between legal and non-legal terms was obvious by the values of any of the indices. This can be explained as resulting from the general definition of ‘legal term’ adopted by the researcher. Statistical indices identify discriminating terms used to index texts, which may or may not be domain terms. So instead of identifying legal terms, statistical indices were used to identify 22 ‘empty terms’, or terms that add little meaning to the text such as ‘book’, ‘chapter’. Terms containing empty terms were removed and the remaining 118,000 terms were considered the legal terms. Legal terms appearing in titles of codes were named fundamental legal terms and identified in a separate sub-list to be used in relation extraction.

The next step after identifying concepts is extracting relations. Four different methods were used to identify relations between terms:

- Syntactical analysis by comparing syntactical roles of terms (e.g. subject-verb, verb-object) and assigning a relation to terms with similar roles.
- Coordination relation analysis in which fundamental legal terms separated by the words ‘and’ and ‘or’ are deemed to be related. Manual validation of the outcome of this method was performed to ensure meaningful results.
- Statistical analysis of the terms’ context words using a cosine similarity measure to compare the vectors of context words for a pair of legal terms and assigning a relation if they share at least 80% of their context.
- Pattern matching in which it is assumed that 2 terms are related if one term is included in the other term (e.g. ‘contract’ and ‘breach of contract’).

The four methods for relation extraction only identify that a relation exists between a two legal terms, however they do not determine the type of this relation. In order to achieve this, the following 5 relation types were defined:

- Is-a-legal-sort-of
- Is-a-general-sort-of
- Is-a-component-of
- Is-related-to
- Is-another-sense-of

The conceptualization stage of ontology building is complete when the previously identified related legal terms are labeled with one of the 5 relation types. The researcher does not detail why the five relation types reported were the ones chosen as labels for the relations existing between the legal terms. The relation types are ambiguous and seem to overlap. For instance, the relation type ‘is-related-to’ seems very ambiguous about the kind of relation existing between two legal terms which have already been identified as being related. Also, the distinction between ‘is-a-sort-of’ and ‘is-another-sense-of’ is not clear. In addition, the researcher did not describe the process of assigning relation types to the sets of related legal terms. If this process was done manually, then the amount of time and effort required to complete this process is expected to be enormous considering the number of the related legal terms (approximately 104,000 terms with at least one relation to another term, approximately 18,000 term with more that one relation to other terms).

The method of identifying legal terms or concepts in this research could be described as extraction by reduction – gradually reducing an initial list of approximately 500,000 terms identified by a parser into a final list of approximately 118,000 legal terms. It is prudent in this case to manually validate both the final list of legal terms and the list of excluded terms in order to ensure that no non-legal terms have been included in the final list and no legal terms have been excluded from the list. Manual validation of course requires lots of time and effort.

Saias and Quaresma (2004) use NLP not only to perform the conceptualization stage of ontology building, but also to automatically develop an ontology from legal documents and merge it into an existent top-level ontology. The top-level ontology is a domain ontology for the Portuguese Attorney General's office, which is already merged with a general top-level ontology expressing general concepts such as concrete object, abstract object, etc. The ontology is integrated into a logic programming framework in order to enable user interaction with the ontology and the legal documents.

Extraction of entities or concepts starts with syntactical analysis of the documents using the shallow parser Palavras. Since the output of Palavras is in VISL (Visual Interactive Syntax Learning), a translator is required to convert it into XML and Prolog. Partial semantic analysis follows in which the output of the syntactical analysis is used to create semantic representations of the sentences of the legal documents by building a discourse representation structure (DRS) for each sentence. From the DRS's, concepts can be extracted thus creating classes that are subclasses of the top class 'Entity'. From the syntactical analysis, properties of the classes can be inferred: adjectives are properties of nouns and direct objects are properties of their transitive verbs.

For the identification of relations, words with similar subcategorization patterns (modify the same words, or are modified by the same words) are assumed to be related. Subcategorization patterns for the identified entities are extracted from the documents and clustered in order to detect words with similar subcategorization patterns. This method can be used to identify hierarchical relations (subclass of) and semantic relations between entities, although in the latter case the type of relation is not identified.

By identifying classes, properties of classes, and the hierarchy of the classes and by determining which classes are related to which, an ontology of the legal documents is created. The next step is to merge the document ontology with the top-level legal ontology mentioned above:

- If a class with the same name exists in the top-level ontology, then both classes are merged.
- If a class with the same name does not exist in the top-level ontology, a similar class based on the semantic compatibility is chosen, and the class from the document ontology is merged into the top-level ontology as a subclass to this class.

An important step for enabling user interaction with the ontology is enriching the legal documents with instances from the developed ontology. This step is performed by translating the ontology expressed in OWL (Ontology Web Language) into Prolog by creating a Prolog term for the different classes, subclasses, and class properties. Then, Prolog rules are encoded that utilize the Prolog-translated ontology to make inferences from the DRS of a document's sentences. This process of inferring instances from semantic representation is called pragmatic interpretation. The outcome of the pragmatic interpretation of the sentences of a certain document is translated back into OWL and stored as specific instances of the ontology in the document.

For user interaction management, a logic programming framework is used. The knowledge base is built by translating the ontology and ontology instances in the documents into SQL tables, and every OWL class instance is given an ISCO

(Information System Construction Language) class definition. ISCO is capable of linking between Prolog terms and databases. EVOLP (Evolving Logic Programming) is used to model user attitudes (intentions and beliefs) and represent actions associated with user queries. A query presented to the system is analyzed syntactically, semantically and pragmatically into an ISCO query using the same techniques described above. Then the knowledge base is accessed to retrieve the most likely entities constrained by the query. The researchers note that the system's capability to deal with queries is limited to restricted queries concerning a specific domain. At present, the system is unable to perform pragmatic interpretation on unrestricted queries.

Although very similar to the work of Lame (2004), since both are concerned with the construction of legal ontologies dedicated to information retrieval, several differences are notable. While Lame's work only concentrated on the extraction of concepts and their existing relations from legal texts, this research aimed at building the legal ontology from the documents, merging it into an existing top-level ontology and facilitating user interaction with the resulting ontology. Also, Lame (2004) uses four different methods to extract relations compared to Saias and Quaresma's (2004) single method. However it is apparent that the latter is more focused on producing a better structured, hierarchical network. This is evidenced in attempting to extract the 'subclass of' relation between entities and in merging the generated document ontology into the existing domain ontology (which in turn is merged with an existing general top-level ontology). Examination of the results of relation extraction in Lame (2004) reveals a flatter structure since about 14,000 terms from a total of 118,000 are not related to any terms and approximately 86,000 terms are related to only one term.

Table 1 compares between the three ontologies that have been previously reviewed: two legal ontologies and one engineering ontology. NLP techniques were not used in developing the engineering ontology. On the other hand, application of NLP techniques to develop the legal ontologies was possible because their domains were expressed in legal texts. The definition of 'concept' varied: in 2 ontologies concepts were defined as noun phrases, while the third ontology identified concepts as both noun and verb phrases. The object-oriented approach in Saias and Quaresma (2004) is evident. The extracted concepts are stored as classes, and attributes for classes are also identified. In Lame (2004) and in El-Diraby and Kashif (2005), there was no attempt to identify concept properties.

Table 1: Comparison of three ontologies

Feature	Lame (2004)	Saias and Quaresma (2004)	El-Diraby and Kashif (2005)
Domain	French Legal codes	Portuguese Attorney General's office documents	Highway design and construction
Type of knowledge source	Textual	Textual	Textual and expert interviews
Use of NLP techniques	Yes	Yes	No
Definition of Concepts	NP	NP and VP	NP
Relations	Set of 5 predefined relation types	No identification of relation types	Verb phrases – undefined list
Attributes	No	Yes	No
Axioms	No	No	Yes

Another important comparison feature is the relation types used in each ontology. Lame (2004) chose to assign relations from a fixed set made up of 5 relation types while Saias and Quaresma (2004) aimed at identifying related concepts without identifying the relation type. El-Diraby and Kashif (2005) allowed a theoretically infinite set of relation types, expressed semantically and dictated by the subject-verb-object relation existing between nouns and verbs. Consequently, relations such as 'influences', 'defines', 'controls' and 'provides' are typically encountered in HiOnto. This method also reflects an object-oriented approach, in which concepts are identified by nouns and relations are identified by verbs (Gómez-Pérez 1998).

Chapter Summary

Knowledge is the most important asset in today's business world and the effective management of knowledge is an important characteristic of successful construction firms. Recent research in KM has attempted to utilize the semantic relations that exist in specific domains to develop KMS's. However the identification of such relations is not an easy task. In the legal field, NLP tools were used to assist in the task of knowledge acquisition.

Contract administration for construction projects is an engineering function that overlaps with the legal domain in specific applications. For example, legal rules and principles play an important role in claims preparation and analysis (Cobb and Diekmann 1986) which is an important function in contract administration. An important aspect in both of these domains is that they rely heavily on textual material expressed in human language: legal references and judicial opinions in the legal domain; contract conditions, specifications, correspondences, etc. in contract administration. It can be said that contract documents express consensual knowledge since they contain the conditions and terms agreed upon between the parties at the time of contract conclusion. Accordingly, these domains are suitable for applying natural language processing techniques to assist in developing knowledge management systems. It is hypothesized that the advancements achieved in the use of NLP techniques to build legal ontologies can be emulated and further developed in the field of construction contract administration.

Chapter 3. Methodology

Overview

System development was based on an object-oriented approach for the identification of concept and concept relations. Concept sets are sets made up of 3 components: the Active Concept, the Relation and the Passive Concept. Concept sets are extracted according to a certain set of rules from the parsed structure of sentences. Parsing was performed by the shallow parser Sundance that is capable of performing certain important functions such as segmenting sentence clauses into various types of phrases and assigning syntactic roles to the phrases. A program in C++ was developed that utilizes Sundance as an application programming interface (API). In order for the program to understand the input file, the input file has to be prepared in a certain format. The program passes sentences to the shallow parser and receives the parsed sentence structure. From this structure the program extract all possible concept sets according to pre-defined rules and the extracted concept sets are stored in a repository.

Concept Sets

The objective of the research is to develop a system for extracting concepts and concept relations from contract documents and evaluate how effective the system is in performing this task. The comparison between the three ontologies reviewed in the previous chapter reveals that there are different approaches regarding the definition of what a concept is (noun phrase, noun and verb phrases), how relations are identified (defined list of relation types, undefined relation types) and how relations are assigned to concepts (manual, automatic). In this research, the object-oriented approach described by Gómez-Pérez (1998) will be used, in which concepts are identified by nouns and relations identified by verbs. Moreover, instead of choosing from a list of predetermined relations, an approach similar to the one used in HiOnto by El-Diraby and Kashif (2005) regarding the relation types will be employed, in which relations are dictated by the subject-verb-object relation found in the text, thereby producing a theoretically infinite set of relation types. Accordingly, a 'concept set' is a set made up of the following three components:

1. Active Concept: the concept doing the action, or the subject of the relation.
2. Relation: the action, or the verb of the relation.
3. Passive Concept: the concept that is being acted upon, or the object of the relation.

For example, suppose the following sentence:

> The Contract Documents form the Contract for Construction.

The Active Concept is the 'Contract Documents', the Relation is 'form' and the Passive Concept is the 'Contract for Construction'. Unlike previous research, concept sets will be extracted either automatically or semi-automatically using NLP techniques, namely the shallow parser Sundance.

Sentence UNDerstanding ANd Concept Extraction: SUNDANCE

Sundance is a natural language parser developed by the School of Computing at the University of Utah (Riloff and Phillips 2004). Sundance is a shallow parser because sentences processed by Sundance are parsed into a two-level structure. The first level is the clause level in which the sentence is broken down into either a single clause or multiple clauses depending on the sentence structure. The second level is the constituent level in which words in each clause are grouped into either noun phrases (NP), verb phrases (VP), prepositional phrases (PP) or adjective phrases (ADJP). In some cases, an orphaned word can be found in this level if the parser is unable to assign it to an adjacent phrase.

Figure 2 gives a sample parse produced by Sundance and the conceptual form of the sentence demonstrating its final two-level structure. As shown in the figure, the constituents at the second level are labeled by their types (NP Segment, VP Segment, PP Segment) and their roles, if any, are identified (SUBJ, DOBJ, and ACTIVE_VERB). By having this simple structure, a computer program can easily access constituents at either of the two levels by looping through the sentence; a single loop will access the clause level, two nested loops will access the constituent level. For example, Clause B can be accessed in the second pass of a single loop, while the PP ‘by the Contract Documents’ can be accessed using a double nested loop, in the second pass of the first loop and the third pass of the second loop.

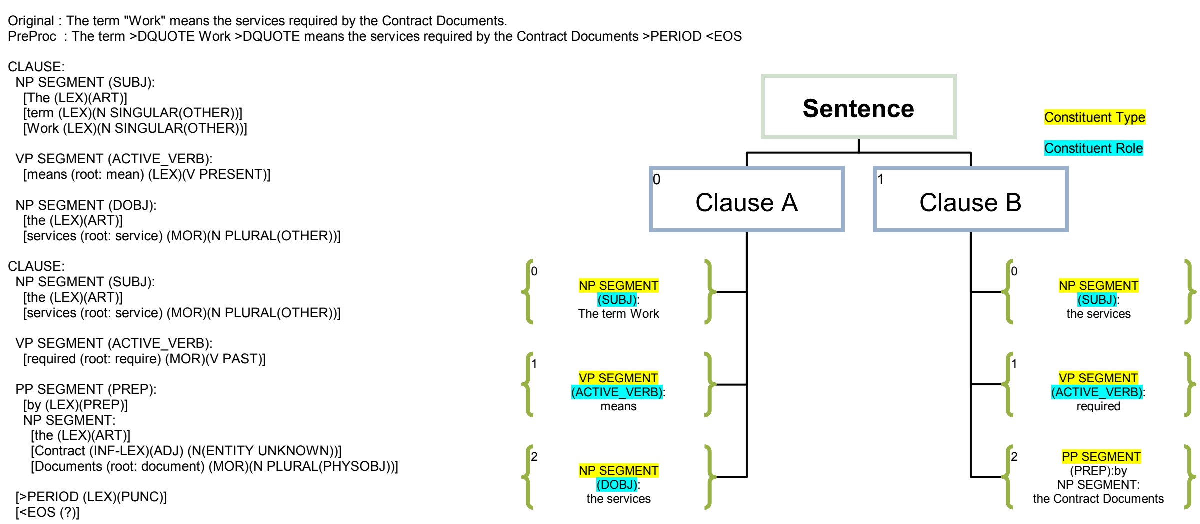

Figure 2: Sample Sundance parse

Sundance can be operated in 3 modes:

- Word tagging mode: Performs part-of-speech tagging either by looking up the words in Sundance's dictionary, or through morphological analysis. A set of 20 part-of-speech tags is used.
- Segmentation mode: Executes the actual parsing of the sentence.
- Extraction mode: Used for extracting specific information from the text. This mode requires initial training to identify the information required for extraction.

The different modes are interdependent; the segmentation mode depends on the word tagging mode and the extraction mode depends on the segmentation mode. In this research the segmentation mode, and consequently the word tagging mode, was used for parsing the text to enable identification of concept sets.

The following is a brief description of some of the important steps taken by Sundance to process a sentence:

- Preprocessing: Includes tokenizing, identification of common phrases to be treated as one word, expanding contractions (e.g. changing 'haven't' into 'have not'), separating punctuation marks from words.
- Sentence segmentation: identification of sentence boundaries.
- Part-of-speech tagging: Initially, a word is tagged with all possible POS tags either through dictionary lookup or morphological analysis. A word therefore can have a single tag, multiple tags, or the tag 'unknown' (if both the dictionary and the morphological analysis can not come up with a tag). Identifying the correct tag actually occurs by using heuristics during syntactic segmentation in which the successive segmenters can eliminate some of the

tags, choose a specific tag, or even override a tag assigned by the part-of-speech tagger.

- Syntactic segmentation: Identifying NPs, VPs, PPs and ADJPs using successive segmenters. This step starts with NP and PP segmenters because these segmenters have the most reliable output which can be used to constrain the output of the other segmenters. The VP segmenter follows in a less conservative manner, since it depends on the output of its preceding conservative and reliable segmenters. Finally, whatever remains of the sentence is passed on to the ADJP segmenter. Phrases identified by each segmenter are labeled with their corresponding constituent type: NP Segment, VP Segment, PP Segment, etc.
- Clause segmentation: Segmenting a single sentence into clauses, in which clauses are identified by either relative pronouns or by multiple VPs.
- Syntactic role assignment: From the previous step, each clause will have only one VP. In this step, Sundance identifies the NPs that are the subject and the direct object of the VP in each clause of the sentence and labels them with the corresponding constituent role: SUBJ, DOBJ. Of course, not all constituents will have a constituent role assigned to them.

In addition to the above steps, Sundance performs other tasks such as named entity recognition, subject inference and even tries to resolve relative pronouns to complete the syntactic role assignment.

It was realized from the onset of the project that the accuracy of the system will be highly sensitive and highly dependant on the accuracy of the NLP tool that is used.

Accordingly the choice of the shallow parser was an important decision. The ability of a shallow parser to perform syntactic segmentation and assign syntactic roles to NPs was an important factor in the choice of the shallow parsing tool that will be utilized by the system. As explained above, Sundance has the ability to perform these essential tasks. In addition, previous research in the field of textual case-based reasoning utilized Sundance mainly in the information extraction mode to index legal cases and develop case representations that can be compared to identify similarities between legal cases and even predict outcomes of the cases (Bruninghaus and Ashley 2005, Bruninghaus and Ashley 2001). Sundance has been described as a robust state-of-the art parser by the researchers and was considered a useful resource for their work. Based on the above, Sundance was chosen as the NLP tool for the system.

Concept Relation Identification using Shallow Parsing – the CRISP Technique

The next step after deciding on Sundance was to develop a computer program for identifying and extracting concept sets from contract documents. Figure 3 illustrates how the proposed system works.

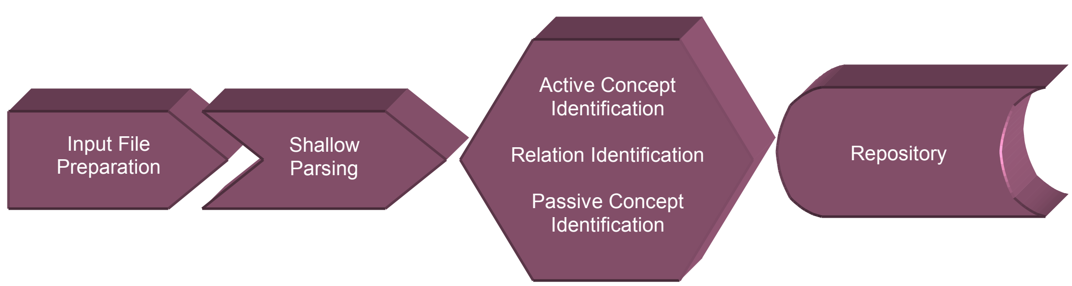

Figure 3: Main components of the CRISP technique

In order for the program to successfully analyze a document presented to it, the document must be initially prepared to a standard format that can be read by the program. Input file preparation can be divided into two steps:

- Basic Preparation: Basic preparation defines the structure of the document to the program. Contract documents are usually divided into sections, subsections, clauses, sub-clauses, etc. For information retrieval purposes and in order to facilitate evaluation of the system, it is important to tag each concept set extracted by the program with the section number it was extracted from. Section numbers in the input file are bound by angle brackets to identify to the program that the sentences following the section number are part of that specific section, and consequently all concept sets extracted from these sentences are tagged with the appropriate section number. In addition, all sections of the document are ended with the dummy sentence 'Clause_End.' to identify to the program the boundaries of each section. Section titles, if available, are not included in the input file. Figure 4 shows a sample of what an input file would look like after basic preparation is complete.

<1.1.3>
The term "Work" means the construction and services required by the Contract Documents, whether completed or partially completed, and includes all other labor, materials, equipment and services provided or to be provided by the Contractor to fulfill the Contractor's obligations. The Work may constitute the whole or a part of the Project. Clause_End.
<1.1.4>
The Project is the total construction of which the Work performed under the Contract Documents may be the whole or a part and which may include construction by the Owner or by separate contractors. Clause_End.
<1.1.5>
The Drawings are the graphic and pictorial portions of the Contract Documents showing the design, location and dimensions of the Work, generally including plans, elevations, sections, details, schedules and diagrams. Clause_End.
<1.1.6>
The Specifications are that portion of the Contract Documents consisting of the written requirements for materials, equipment, systems, standards and workmanship for the Work, and performance of related services. Clause_End.

Figure 4: Sample input file after basic preparation

- Advanced Preparation: Advanced preparation can be performed in order to improve the performance of the shallow parser. Further details on why advanced file preparation is required and how it is performed are given in the following chapter.

A computer program was developed in C++ to manipulate the various components of the system for the purpose of extracting the concept sets. The main program uses Sundance as an API in order to access Sundance's libraries and utilize its parsing capabilities, therefore Sundance must be initialized at the start of the program and its source code and libraries must be defined in the program's makefile. The main program reads from the input file. If a section number is encountered, a specific program variable is updated with the section number. If the section body is encountered it is passed to Sundance. The program depends on Sundance for sentence boundary identification. Sundance extracts a sentence from the text, parses it, and returns it to the main program as an object of type Sentence having the two-level structure explained above. The main program then sets out to utilize this Sentence object to extract concept sets.

A class named ConceptRelation was developed that takes two parameters: the Sentence object and the section number to which the original sentence belongs. The purpose for developing this class is to extract all possible concept sets from a single Sentence object. Accordingly, an object of ConceptRelation type is created for each sentence parsed by Sundance. Once a ConceptRelation object is created, the object analyzes the parsed sentence to extract the Active Concept, Relation and Passive Concept in each clause of the sentence according to the following rules:

- Active Concept: The segments under each clause are checked for a constituent with role 'SUBJ'. If the subject segment is identified, it is extracted and considered the Active Concept component of the concept set. Constituents following the subject segment are checked to identify PPs that follow the subject. PPs following the subject are considered to qualify the Active Concept and are therefore extracted along with the subject segment. Figure 5 illustrates this process as a flowchart.

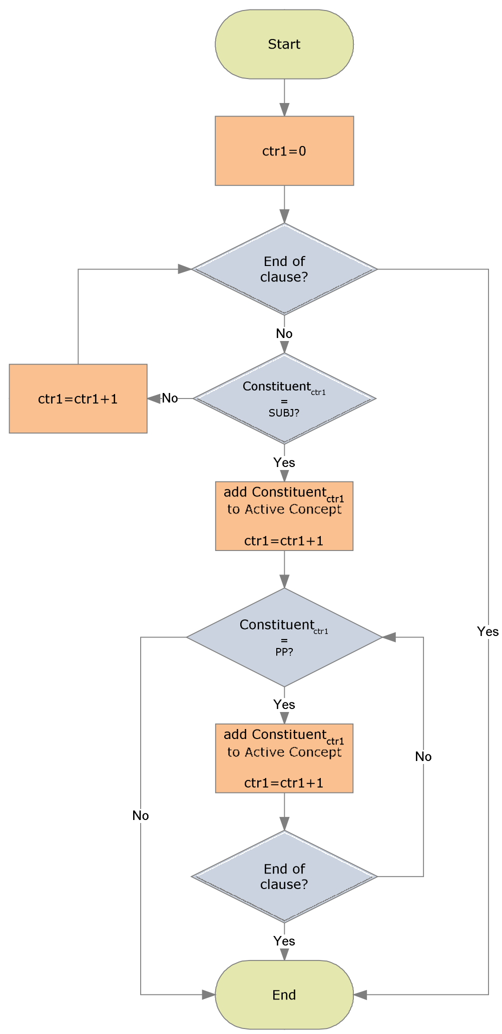

Figure 5: Active Concept extraction

- Relation: As previously mentioned, sentences are segmented into clauses based on the existence of multiple VPs, and consequently each clause will have only one VP. The segment under each clause with constituent type 'VP Segment' is extracted and considered the Relation component of the concept set.

  Passive Concept: The segments under each clause are checked for a constituent with role 'DOBJ'. If the object segment is identified, it is extracted and considered the Passive Concept component of the concept set. Constituents following the object segment are checked to identify PPs that follow the object. PPs following the object are considered to qualify the Passive Concept and are therefore extracted along with the object segment. As an example, Figure 6 shows a shallow parse for a sample sentence.

  After parsing, the NP 'the Contract' is identified as the direct object. If only the direct object is extracted, then the Passive Concept will be limited to 'the Contract' instead of 'the Contract for Construction'. Since the PP 'for Construction' adds meaning to the Passive Concept, it is also extracted and the Passive Concept becomes 'the Contract for Construction'.

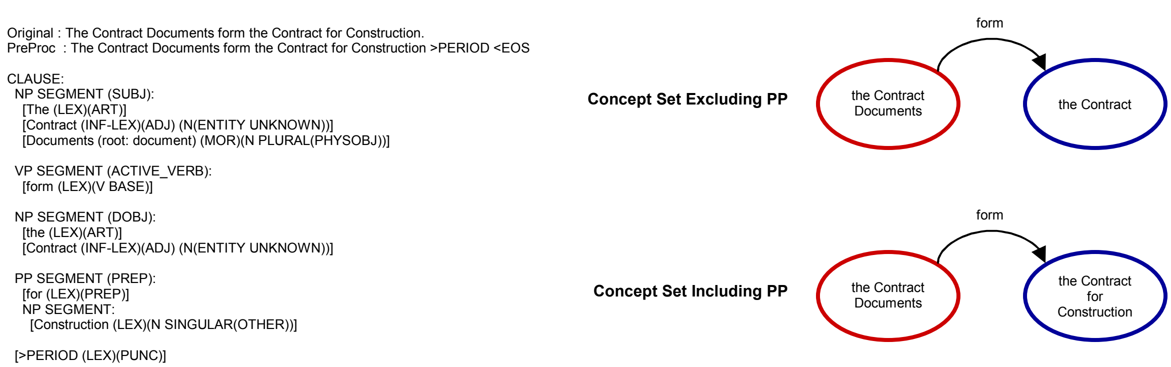

Figure 6: Qualifying concept set components with PPs

In some cases, no direct object is identified in the clause. This especially occurs when the VP is followed by a PP. Consider the following sentence:

The Contract Documents consist of the Agreement between Owner and Contractor.

The VP 'consist' is followed by the PP 'of the Agreement'. In other words the object – 'Agreement' – is an NP forming part of the PP, but since it is not directly one of the clauses constituents, the shallow parser does not label it as 'DOBJ', and accordingly the program cannot extract a Passive Concept for this concept set. To overcome this, in the case where all clause segments are checked and none have a constituent role of 'DOBJ', the constituent type for the segment directly after the VP is checked. If it is a PP, then it is extracted and assumed to be the Passive Concept component of the concept set. In the above example, 'of the Agreement' will be correctly extracted as the Passive Concept. Here, also, constituents following the assumed object segment are considered to qualify the object and are also extracted and added to the Passive Concept component. Figure 7 illustrates this process as a flowchart.

Concept sets extracted from each ConceptRelation object are stored in a data-member of the object. After parsing all sentences in the input file by Sundance and analyzing them by the class ConceptRelation, the main program loops through all ConceptRelation objects and collects all concept sets into one large data structure, the system's repository. This repository will be the subject of the system's evaluation and will also be the main component for utilizing the system for future work. The computer code for class ConceptRelation and the main program is given in Appendix A.

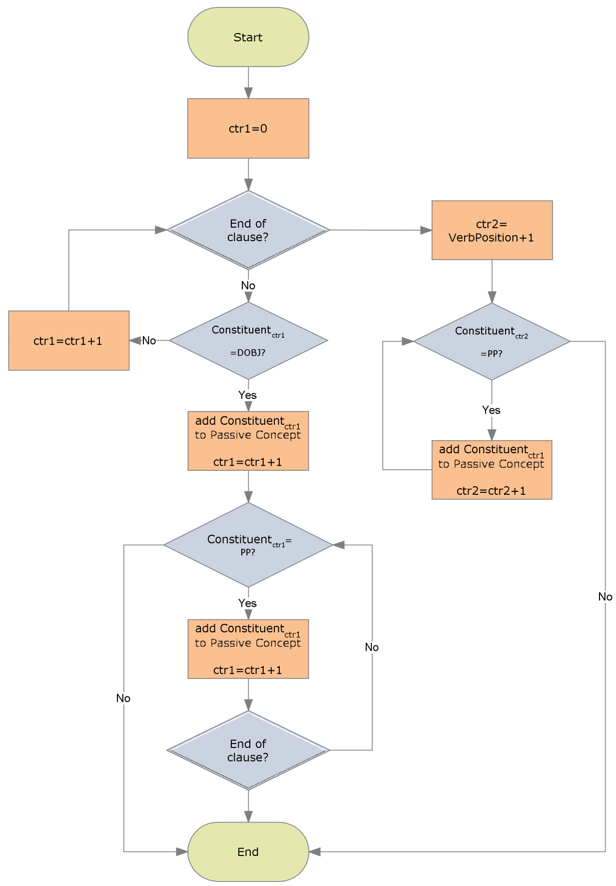

Figure 7: Passive Concept extraction

Chapter Summary

In this chapter, the methodology for developing CRISP was described. A brief description of the shallow parser used by CRISP, Sundance, was first presented. Then the basic preparation steps that must be performed on the input file were detailed. Finally, the rules used by CRISP to extract the Active Concept, Relation and Passive Concept from the parsed structure of the sentence were stated. The following chapter describes the process used to evaluate CRISP.

Chapter 4. System Evaluation

Overview

The American Institute of Architect's General Conditions of the Contract for Construction (AIA Document A201™ – 1997) was used for the evaluation of CRISP. Initial parsing trials from random sentences in the document revealed that certain text features can cause significant parsing errors by the shallow parser. Accordingly, it was decided to prepare two input files, an original input file based on basic preparation only, and a modified input file based on advanced preparation. Evaluation of CRISP was to be performed by comparing the system's output to the results of 7 human evaluators. Since the evaluation of the whole test document by human evaluators was not possible, a sample set of provisions was prepared for the evaluation taking care to avoid any bias in the selection of provisions. Human evaluator results were compiled and compared to the two outputs from CRISP, Output1 from the original input file and Ouput2 from the modified input file. Average agreement between human evaluators was relatively low at 0.410. Output1 achieved an average agreement with human evaluators of 0.185, while Output2 achieved an average agreement with human evaluators of 0.323. A Gold Standard was prepared from the analysis of the human evaluations for calculating precision and recall. The average F-measure score of the human evaluators was 0.758. Output1 achieved an F-measure score of 0.420, while Output2 achieved a score of 0.681.

Test Document

The first step in the evaluation procedure is to choose an input document to be analyzed by CRISP. Construction contract documents usually include a variety of textual documents expressed in natural language (e.g. the Agreement, the General Conditions of the Contract, the Specifications, etc.). It was decided to conduct the test on a standard form of contract, namely the American Institute of Architect's General Conditions of the Contract for Construction (AIA Document A201™ – 1997) for the following reasons:

- Standard forms of contract are published by organizations with many years of experience in the construction industry and therefore represent the industry standard for the contractual agreement between the different project participants.
- Standard forms of contract are widely accepted by all project participants and widely used in the construction industry. Accordingly, a meaningful analysis of standard forms can be beneficial to all projects that use the standard form.
- Because they undergo numerous revisions, standard forms are error free.
- AIA documents have evolved over almost 115 years through numerous editions to become benchmark documents expressing the contractual relationships between construction parties (The American Institute of Architects 2008). AIA Document A201™ – 1997 is the 15th edition of the document, with the first edition dating back to 1888.

- AIA Document A201™ – 1997 is the most common standard form of contract used by prime contractors in building projects in the United States (The American Institute of Architects 1999).
- AIA Document A201™ – 1997 is considered among other AIA documents as the keystone document for the traditional design-bid-build project delivery method.

Characteristics of the Input File

AIA Document A201™ – 1997 is made up of 14 articles, each article relating to an important topic in construction projects such as General Provisions, Changes in the Work, Payments and Completion, etc. Articles are made up of sections. Sections may be divided into subsections, and subsections may be divided into sub-subsections. All in all, AIA Document A201™ – 1997 is made up of 264 provisions, a total of 19,679 words, with an average of 75 words per provision.

An important feature found in standard forms of contract in general and in AIA Document A201™ – 1997 in specific is the presence of enumerations and lists in the provisions' text. Figure 8 gives a sample enumeration and a sample list in a sentence from the text of AIA Document A201™ – 1997.

Sample enumeration:
The Owner shall furnish surveys describing physical characteristics, legal limitations and utility locations for the site of the Project, and a legal description of the site.

Sample list:
A Modification is (1) a written amendment to the Contract signed by both parties, (2) a Change Order, (3) a Construction Change Directive or (4) a written order for a minor change in the Work issued by the Architect.

Figure 8: Sample enumeration and list in the text

Initial testing of Sundance with sentences containing enumerations or lists showed that the accuracy of the parser in segmenting a sentence into clauses decreases, and accordingly syntactic role assignment is affected. In addition, enumerations and lists produce multiple Active Concepts, Relations and/or Passive Concepts in a single sentence clause. Therefore, even if the clause segmentation was correct, Sundance will identify only one NP with role SUBJ and only one NP with role DOBJ per clause, ignoring the multiplicity. Since the performance of the whole system depends largely on the accurate assignment of syntactic roles, and due to the abundant use of lists and enumerations in the input document, this text feature had a critical effect on the performance of CRISP.

Based on this finding, it was decided to prepare a modified input file and evaluate CRISP on both the original input file (prepared using the basic preparation steps described in Chapter 3) and the modified input file (prepared according to the advanced preparation steps). Advanced preparation, which also includes the basic preparation steps required in order to define the structure of the document to the program, is used to resolve the issues created by enumerations and lists in the text. In advanced preparation, sentences containing enumerations and lists are elaborated to facilitate accurate parsing of the sentences. Basically the same process is used to resolve both enumerations and lists:

- For enumerations, the sentence is broken down into separate sentences, each containing a component of the enumeration.
- For lists, the numbering of the list is removed and the sentence is repeated for each individual member of the list.

This process is a systematic manual process that does not look at the grammatical structure of the sentence. In other words, the resulting sentences after the resolution of enumerations or lists can contain grammatical errors. No effort is made to correct such errors; CRISP relies on the abilities of Sundance to parse ungrammatical sentences. Figure 9 shows the final form of the previous samples after advanced preparation.

Sample enumeration after resolution:
The Owner shall furnish surveys describing physical characteristics for the site of the Project.
The Owner shall furnish surveys describing legal limitations for the site of the Project.
The Owner shall furnish surveys describing utility locations for the site of the Project.
The Owner shall furnish a legal description of the site.

Sample list after resolution:
A Modification is a written amendment to the Contract signed by both parties.
A Modification is a Change Order.
A Modification is a Construction Change Directive.
A Modification is a written order for a minor change in the Work issued by the Architect.

Figure 9: Final form after enumeration/list resolution

Advanced preparation can be a tedious process, especially when sentences contain multiple enumerations or combined enumerations and lists that result in numerous possible combinations. The purpose of preparing a modified input file was to evaluate CRISP's performance independent of the accuracy of the shallow parser and compare how the system will perform with and without the parser's handicap.

Evaluation Process

A simple evaluation process was used. The concept sets extracted by CRISP from the two input files were compared to the concept sets extracted by human evaluators for precision, recall and agreement with human evaluators. Because it is unpractical to have human evaluators evaluate all 264 provisions of AIA Document A201™ – 1997, an evaluation set of provisions had to be developed. Initially, it was decided to present the human evaluators with 10 provisions. To avoid biased selection, a random number generator was used to select 15 of the 264 provisions. Of these, 10 provisions were selected according to the following criteria:

- To avoid any bias, either the full provisions is selected or it is excluded from the evaluation set; partial provisions were not used.
- The average words per provision over the whole document is 75, accordingly the evaluation set must have an equivalent average words per provision.
- The selection of provisions must not be from a certain part of the document. The chosen provisions must be distributed over the whole body of the document.
- The provisions must include the common features found in standard forms of contract such as enumerations and lists.

In addition to these criteria, it was important that the evaluation should not be a heavy burden on the human evaluators in order to get comprehensive and accurate results. A review of the 10 provision evaluation set resulted in a decision to reduce the evaluation set to only 6 provisions. The final evaluation set conformed to the required criteria; 6 full provisions distributed over the whole document, with an average of 76 words per provision.

Both original and modified input files were prepared for the evaluation set and processed by the program. The input files are given in Appendix B. The output of CRISP was prepared in a spreadsheet format to facilitate comparison with human output from the evaluation session. Appendix C presents the output of CRISP for both input files.

A two hour evaluation session was organized for the 7 human evaluators. All evaluators were graduate students in civil engineering with practical experience in the construction industry. All of them had previously completed at least one course in construction contracts, the majority completing two or more. A short 10 minute presentation was given to outline the objective of the exercise, explain what concept sets are and demonstrate the extraction of Active Concepts, Relations and Passive Concepts from simple sentences. The evaluation exercise (Appendix D) was then presented to the evaluators and the evaluators were instructed to try to be as comprehensive as possible in identifying the concept sets. The results of the human evaluators were compiled and prepared in a spreadsheet format.

Preliminary Results

In terms of the number of concept sets extracted by each evaluator and by CRISP for both input files, Figure 10 shows that the number of concept sets extracted by the system from the original input file (Output1) for every subsection was generally on the lower boundary of the numbers extracted by the human evaluators, while the number of concept sets extracted from the modified input file (Output2) was generally around the average number extracted by human evaluators. Although these numbers show that CRISP extracted concept sets within the expected limits, the number of concept sets does not reflect the performance of the system. For an accurate evaluation of performance, kappa scores and precision and recall values were calculated for both the human evaluators and CRISP's outputs.

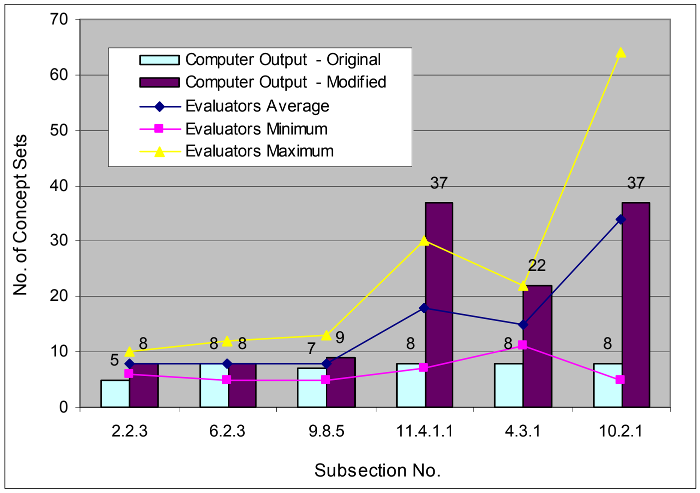

Figure 10: Number of extracted concept sets

Kappa Scores

Kappa measures the pairwise agreement between two human evaluators, or a human evaluator and the system's output after adjusting for chance agreement (Jurafsky and Martin 2000). Kappa is calculated according to the following equation:

$$k = \frac{P(A) - P(E)}{1 - P(E)}$$

where P(A) is the probability of agreement and P(E) is the probability of chance agreement.

The result of each human evaluator was compared in a pairwise comparison with the results of the other human evaluators and the results of CRISP. All in all, there were 35 comparisons; 21 human-human comparisons, and 14 human-computer comparisons (7 with Output1 and 7 with Output2). In a pairwise comparison, kappa is calculated between the concept sets extracted by the two evaluators for each subsection. Table 2 gives the concept sets extracted by Evaluator C and Evaluator D for subsection 2.2.3. Highlighted concept sets are the sets agreed on by both evaluators. In total, the two evaluators extracted 16 concepts. Considering the concept sets agreed on by the two evaluators, the union of the concept sets (i.e. the actual total) is 9 and the intersection is 7. Assuming equal probability for picking a concept set:

$$P(E) = \left(\frac{1}{9}\right)^2 \qquad P(A) = \frac{7}{9} \qquad k = \frac{(7/9) - (1/9)^2}{1 - (1/9)^2} = 0.775$$

Table 2: Comparison between results of Evaluators C & D for subsection 2.2.3

Ser.	Evaluator C			Evaluator D		
	Active Concept	Relation	Passive Concept	Active Concept	Relation	Passive Concept
1	the Owner	shall furnish	surveys	Owner	furnish	surveys
2	surveys	describing	physical characteristics for the site of the project	Owner	furnish	legal description
3	surveys	describing	legal limitations for the site of the project	Contractor	rely	Information
4	surveys	describing	utility locations for the site of the project	Contractor	exercise	Precautions
5	the Owner	shall furnish	a legal description of the site	Surveys	describing	physical characteristics
6	the Contractor	rely	the accuracy of the information	Surveys	describing	legal limitations
7	information	furnished by	the Owner	Surveys	describing	utility locations
8	the Contractor	shall exercise	proper precautions			
9	precautions	relating to	the safe performance			
Total	9			7		

An average kappa over the 6 subsections is calculated representing the average agreement between the two evaluators over the complete evaluation set. Finally an average kappa is calculated over the 21 human-human comparisons which is compared to the average kappa of the 7 human-Output1 comparisons and the 7 human-Output2 comparisons. Table 3 gives the overall results for the kappa scores and Appendix E lists the results of the pairwise comparisons between the evaluators.

Table 3: Kappa scores

Evaluators	Evaluator A	Evaluator B	Evaluator C	Evaluator D	Evaluator E	Evaluator F	Evaluator G	Average Evaluators
Evaluator A	--	0.280	0.403	0.394	0.353	0.289	0.433	0.359
Evaluator B	0.280	--	0.469	0.289	0.282	0.302	0.482	0.351
Evaluator C	0.403	0.469	--	0.485	0.382	0.479	0.759	0.496
Evaluator D	0.394	0.289	0.485	--	0.366	0.370	0.469	0.396
Evaluator E	0.353	0.282	0.382	0.366	--	0.369	0.382	0.356
Evaluator F	0.289	0.302	0.479	0.370	0.369	--	0.566	0.396
Evaluator G	0.433	0.482	0.759	0.469	0.382	0.566	--	0.515
Average Evaluators	0.359	0.351	0.496	0.396	0.356	0.396	0.515	0.410
Output1	0.098	0.203	0.230	0.186	0.194	0.172	0.213	0.185
Output2	0.343	0.307	0.401	0.294	0.304	0.239	0.376	0.323

The average agreement between human evaluators is 0.410. CRISP achieved an average agreement with human evaluators of 0.185 with Output1. This number increased to 0.323 with Output2, approximately 79% of the result achieved by human evaluators. Figure 11 shows the average agreement achieved by each evaluator with the other human evaluators and the system's agreement with each evaluator.

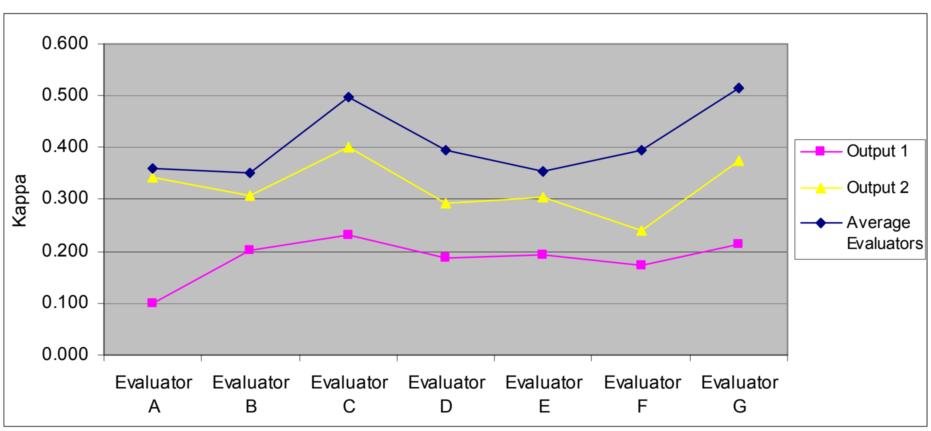

Figure 11: Average kappa scores

The figure demonstrates how both computer outputs followed the general agreement trend of the human evaluators. For example, Evaluators C and G achieved the highest average kappa scores between all human evaluators. Likewise, both computer outputs achieved their highest kappa scores in comparisons with Evaluators C and G.

Precision and Recall

Precision, or accuracy, measures the percentage of correct concept sets from the total number of concept sets extracted by the evaluator/system, while recall, or coverage, measures the percentage of correct concept sets extracted by the evaluator/system from the total number of correct concept sets. In order to measure precision and recall, the correct concept sets– the Gold Standard– must be determined from the results of the human evaluation. A total of 640 concept sets were identified by the evaluators. The number of repetitions of each concept set extracted by the evaluators was identified. Repetition values ranged from a maximum of 7 (for a concept set identified by all the evaluators) to a minimum of 1 (for a concept set identified by only one evaluator). Considering repetition, the actual total number of concept sets was determined at 314. The next step was to determine a threshold of repetitions according to which a concept set is considered correct. A low threshold increases precision and reduces recall, while a high threshold decreases precision and increases recall. In determining the threshold, the kappa scores achieved by the evaluators were considered. High agreement between the evaluators would justify the use of a high threshold. However, since evaluator agreement proved to be relatively low, a threshold that is slightly towards the low end was used. A concept set is considered correct if at least 3 evaluators agreed on the concept set. Concept sets that satisfy this criteria were gathered to develop the Gold Standard. Of the 314 concept sets, 71 concept sets made the Gold Standard. Appendix F lists the Gold Standard for each of the 6 subsections in the evaluation set.

Precision and recall were calculated by comparing the results of the human evaluators and computer outputs with the Gold standard for each subsection in the evaluation set. Table 4 displays the Gold Standard and Output1 for subsection 2.2.3. Highlighted concept sets in Output1 are the correct concept sets as determined by the Gold Standard. Four correct concept sets have been extracted by CRISP. The precision of CRISP is therefore four correct concepts out of a total of 5 extracted concepts:

$$\text{Precision} = \frac{4}{5} = 0.800$$

The recall of CRISP is four correct concepts out of a total of nine Gold Standard concepts:

$$\text{Recall} = \frac{4}{9} = 0.444$$

Precision and recall values are averaged over the six subsections in the evaluation set to determine the average precision and recall of a human evaluator or a computer output. Table 5 gives the average precision and recall values as well as an overall average for all human evaluators, while Appendix G lists detailed precision and recall values for the human evaluators and the system's outputs over the various subsections in the evaluation set.

Table 4: Gold Standard and Output1 for subsection 2.2.3

Ser.	Gold Standard			Computer Output - Original		
	Active Concept	**Relation**	**Passive Concept**	**Active Concept**	**Relation**	**Passive Concept**
1	The Owner	shall furnish	a legal description of the site	The Owner	shall furnish	surveys
2	The Contractor	shall be entitled to rely on	the accuracy of information furnished by the Owner	surveys	describing	physical characteristics
3	The Contractor	shall exercise	proper precautions relating to the safe performance of the Work	The Contractor	shall be entitled to rely	on the accuracy of information
4	Owners	shall furnish	surveys	the accuracy of information	furnished	by the Owner
5	Surveys	describing	physical characteristics	the Owner	shall exercise	proper precautions relating to the safe performance of the Work
6	Surveys	describing	legal limitations			
7	Surveys	describing	utility locations			
8	information	furnished by	Owner			
9	precautions	relating	safe performance			
Total	9			5		

Table 5: Average precision and recall values

Evaluator	Precision	Recall
Evaluator A	0.834	0.615
Evaluator B	0.561	0.858
Evaluator C	0.904	0.929
Evaluator D	0.895	0.588
Evaluator E	0.734	0.633
Evaluator F	0.708	0.712
Evaluator G	0.842	0.978
Average Evaluators	0.783	0.759
Output1	0.476	0.376
Output2	0.698	0.665

F-measure is used to combine precision and recall into one measure (Jurafsky and Martin 2000). F-measure is calculated as follows:

$$\text{F-measure} = \frac{(\beta^2 + 1)PR}{\beta^2 P + R}$$

where P is the precision value and R is the recall value.

β is a parameter that is used to assign relative weights to the precision and recall values. At this stage, precision and recall are considered of equal importance; accordingly the value for β is one.

Figure 12 compares the average F-measure score for all human evaluators and the F-measure scores for Output1 and Output2. Consistent with previous results, a significant improvement in the performance of CRISP was observed when advanced preparation was used to modify the input file.

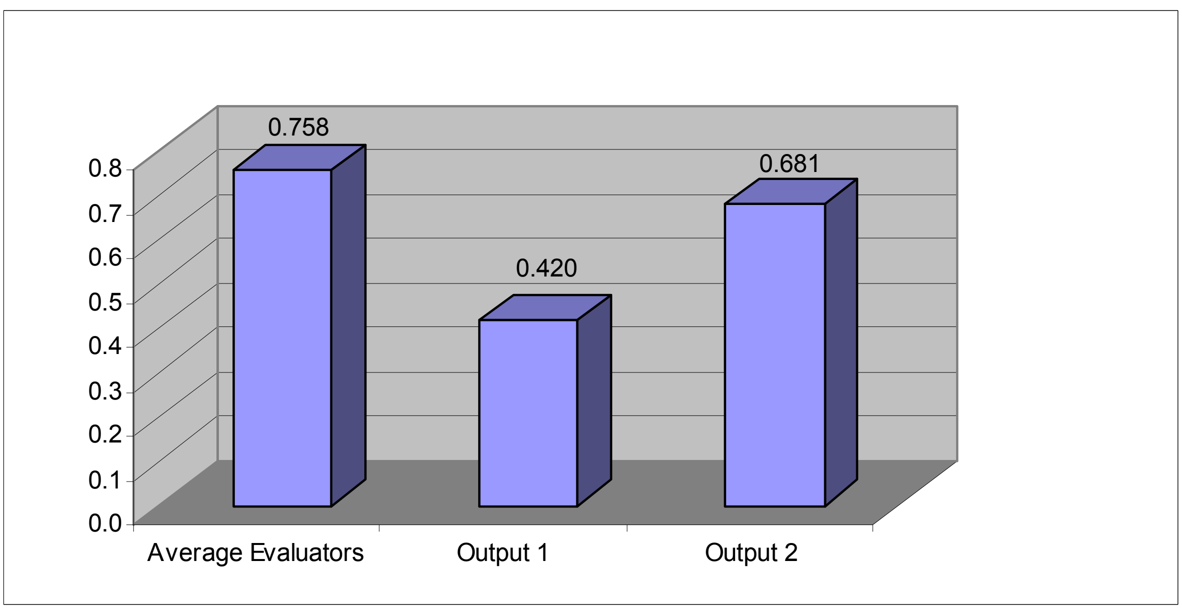

Figure 12: Average F-measure scores

Chapter Summary

A description for the evaluation process for CRISP was given in this chapter. The document used for the evaluation was presented and described and the reasons for its selection were stated. The reasons for the use of advanced preparation were given and the steps of advanced preparation were detailed. The human evaluation process was then described, and finally the results from the comparison between CRISP and the human evaluators were presented in terms of kappa and F-measure scores. The next chapter presents several ideas on how CRISP can be improved in future work and describes a few applications that can benefit from the use of the system.

Chapter 5. Conclusion and Future Recommendations

Conclusions

As expected, the performance of the system was highly dependent on the performance of the shallow parser used. A review of the incorrect concepts extracted by the system revealed two main parsing errors:

- Errors resulting from complex features in the input files, namely enumerations and lists.
- Errors resulting from incorrect syntactic segmentation, namely the incorrect identification of nouns as verbs and the consequent formation of incorrect VPs.

Advanced input file preparation was introduced in an attempt to reduce the effect of the first parsing error. Advanced preparation may be a manual and tedious process. However, the logic behind the idea was to try to evaluate the system's performance independent of Sundance's specific inaccuracies and determine whether or not a little bit of effort in input file preparation will be rewarded with improved performance. Indeed, a significant improvement in performance was observed:

- 75% increase in kappa scores
- 62% increase in F-measure scores

Correction of the second parsing error required significant re-coding of Sundance to modify the heuristics employed by the successive phrase segmenters. Accordingly it was decided to ignore this error (especially since its effects on the results were not as drastic as the first error's effects) thereby absorbing its effects into the performance of the overall system.

Results obtained by the system can be considered encouraging. The standard agreement level between human evaluators was 0.410. The system in its best output achieved an agreement level of 0.323, approximately 80% of the average human standard. Moreover, the system displayed agreement trends that were equivalent to the trends observed from the human evaluations. In terms of precision and recall, the average F-measure of all human evaluators based on the established Gold Standard was 0.758. The best F-measure achieved by the system was 0.681; approximately 90% of the average human standard.

Future Works

From these results, future works on CRISP to improve performance may include the following developments:

- Automating the advanced preparation process: By doing so, the improvement in performance resulting from applying advanced preparation can be achieved while avoiding the manual effort involved. However, automation of the advanced preparation process is in itself a significant NLP problem.
- Testing CRISP with other shallow parsers: As previously discussed, the performance of CRISP is greatly affected by the accuracy of the NLP tool used. It is possible that better system performance can be achieved by using another shallow parser. However in this case, the main program must be specifically adapted for each shallow parser to enable the use of its capabilities.
- Using a deep parser: Table 6 lists the Gold Standard and Output1 for the following sentence:

The Owner shall furnish surveys describing physical characteristics, legal limitations and utility locations for the site of the Project, and a legal description of the site.

Table 6: Gold Standard and Output1 for sample sentence

Ser.	Gold Standard			Computer Output - Original		
	Active Concept	Relation	Passive Concept	Active Concept	Relation	Passive Concept
1	The Owner	shall furnish	a legal description of the site	The Owner	shall furnish	Surveys
2	Owners	shall furnish	surveys	surveys	describing	Physical characteristics
3	Surveys	describing	physical characteristics			
4	Surveys	describing	legal limitations			
5	Surveys	describing	utility locations			

Because of the enumeration, Sundance only identified ‘surveys’ as the DOBJ of ‘shall furnish’ and ignored ‘legal description’. Similarly, ‘physical characteristics’ was identified as the DOBJ of ‘describing’, ignoring ‘legal limitations’ and ‘utility locations’. Figure 13 shows the parse tree generated by a deep parser (Stanford parser) for the same sentence. It is apparent that the output of the deep parser was slightly more accurate. The NP ‘a legal description’ was identified as related to the verb ‘furnish’. However, ‘legal limitations’ and ‘utility locations’ were incorrectly related to the verb ‘furnish’ instead of the verb ‘describing’. The following challenges are anticipated if a deep parser is used:

- Syntactic role assignment: As demonstrated in Figure 13, syntactic roles are not assigned by the deep parser. Syntactic role assignment was one of the most important features used for concept set extraction. It is possible that a deep parser exists that performs syntactic role assignment. However, in the absence of this feature, rules must be defined in the program's code to identify the Active Concept and the Passive Concept (for example, the first NP before the verb and the NPs following the verb until another verb is encountered).

- Indefinite number of levels in the parse structures: The two-level structure of the shallow parser discussed in Chapter 3 simplified development of the computer program. Figure 13 shows that this is not the case in a deep parse. Depending on the sentence, indefinite number of levels can exist that vary throughout the parse which will increase programming difficulty.

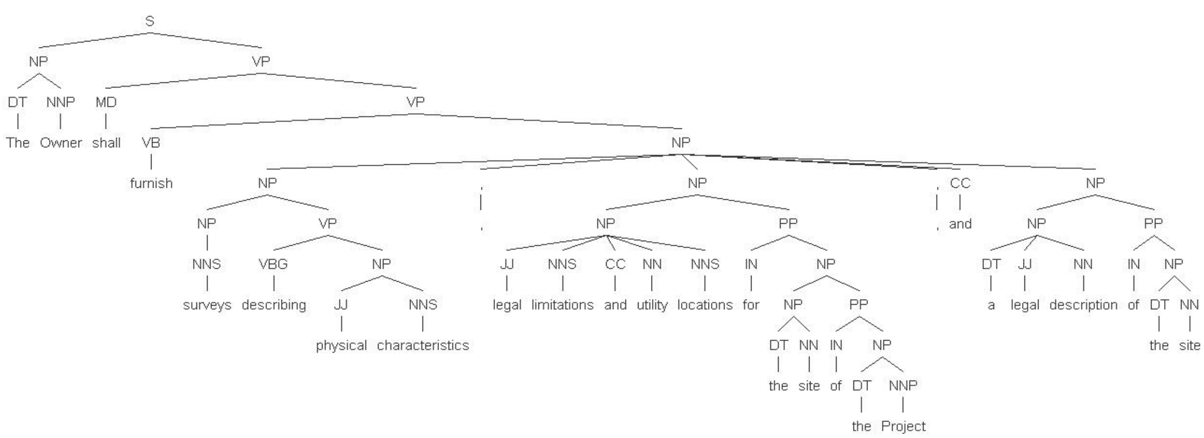

Figure 13: Output of a deep parser

From this example, it can be hypothesized that using a deep parser can partially reduce the detrimental effect of enumerations and lists, providing results that are better than the results of a shallow parser using basic input file preparation, but not as good as the results achieved by using advanced preparation.

Possible Applications for CRISP

Although the purpose for developing the system was to enhance KM techniques in the construction industry, CRISP can be used on any textual documents and is therefore not limited to a specific domain. The proposed system can be utilized for various applications:

- As discussed in Chapter 2, recent research has been focusing on the use of ontologies to facilitate KM. Ontologies not only provide a common foundation for knowledge exchange but can also be used to enable natural language communication with computers. However the potential advantages derived from the use of ontologies have not been fully realized due to the difficulties encountered in building ontologies. The knowledge required for building ontologies is collected from various sources, and in many cases is expressed in natural language in the form of textual documents. CRISP can be used to assist in the difficult and time-consuming process of extracting concept and concept relation from texts, thus assisting in the knowledge acquisition stage of ontology building.
- CRISP can be utilized as a component of an information retrieval (IR) system to enable querying in natural language, regardless of whether or not the IR system is an ontology-based system. Documents in the IR system are tagged with the concept sets extracted by CRISP. A query submitted to the IR system is analyzed by CRISP and a two-component concept set is extracted. The repository of concept sets is accessed and the concept set(s) that best resembles the query is identified by comparing components. The answer to

the query is therefore the missing component that completes the query's concept set, and the document that contains the answer is consequently identified. As an example consider the following question:

Who shall furnish surveys?

Figure 14 shows the result of the shallow parse of the question, the two-component concept set extracted by CRISP from the query and the concept sets for subsection 2.2.3 in Output1. The concept set from Output1 that best resembles the query's concept set is the first concept set. Accordingly, the answer to the query is 'surveys' and the answer is located in subsection 2.2.3.

```
Original : who shall furnish surveys?
PreProc  : who shall furnish surveys >QUESTION <EOS

CLAUSE:
 [who (LEX)(C_M)]
 VP SEGMENT (ACTIVE_VERB):
   [shall (LEX)(AUX)]
   [furnish (LEX)(V BASE)]

 NP SEGMENT (DOBJ):
   [surveys (root: survey) (MOR)(N PLURAL(OTHER))]

 [>QUESTION (LEX)(PUNC)]

 [<EOS (?)]
```

Query

?	shall furnish	surveys

Computer Output1

The Owner	shall furnish	surveys
surveys	describing	physical characteristics
The Contractor	shall be entitled to rely	on the accuracy of information
the accuracy of information	furnished	by the Owner
the Owner	shall exercise	proper precautions relating to the safe performance of the Work

Figure 14: Automated query answering framework

The preceding example was a conceptual description of how CRISP can be used to answer natural language queries. Obviously at the current stage these capabilities are very limited and require further development (for example, using a lexicon to generate various possible forms of the query's concept set in order to enhance the repository search and increase the possibility of finding an answer).

- CRISP can be used to develop UML diagrams of the important concept and concept relations in text documents. In contract documents, this can be used, for example, to illustrate the roles of the major parties of the contract. Figure 15 is a sample UML diagram for the concept sets extracted from subsection 2.2.3 identifying the duties of the Owner and Contractor dictated by this subsection. The diagram presents a visual summarization of the subsection which can be helpful for educational and training purposes.

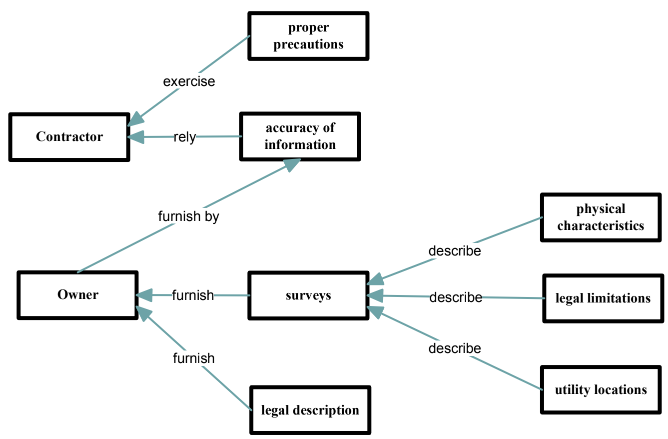

Figure 15: UML diagram for subsection 2.2.3

References Cited

Allen, James (1995). *Natural language understanding*, 2nd Ed., Benjamin/Cummings, California.

American Institute of Architects, Heuer, Charles R., and Goldberg, Howard G. (1999). *Commentary on AIA Document A201-1997*, The American Institute of Architects, Washington D.C.

American Institute of Architects (2008). “History of contract documents.” <http://www.aia.org/docs_history> (March 2008).

Anandha Mala, G.S., and Uma, G.V. (2006). “Automatic construction of object oriented design models [UML Diagrams] from natural language requirements specification.” *Proc., PRICAI 2006: Trends in Artificial Intelligence - 9th Pacific Rim International Conference on Artificial Intelligence*, 4099, 1155-1156.

Asprey, L. (2004). “Information strategies: Are we aligning the business case with enterprise planning?” *Rec. Manage. J.*, 14(1), 7–13.

Borst, W.N. (1997). “Construction of Engineering Ontologies.” *Technical Report, Centre for Telematica and Information Technology*, University of Tweenty Enschede, The Netherlands.

Brüninghaus, S., and Ashley, K. D. (2001). “The role of information extraction for textual CBR.” *Proc. 4th International Conference on Case-Based Reasoning*, Springer-Verlag, Berlin, Germany, 74-89.

Brüninghaus, S. and Ashley, K. (2005). “Reasoning with textual cases.” *Case-Based Reasoning Research and Development*, Springer-Verlag, Berlin, Germany, 137-151.

Carneiro, Alberto (2001). “The role of intelligent resources in knowledge management.” *Journal of knowledge management*, 5(4), 358-367.

Carrillo, P.M., Robinson, H.S., Al-Ghassani, A.M. and Anumba, C.J., (2004). “Knowledge management in UK construction: strategies, resources and barriers.” *Proj Manage J*, 35(1), 46-56.

Chinowsky ,Paul, and Carrillo, Patricia (2007). “Knowledge management to learning organization connection.” *J. Mgmt. in Engrg.*, 23(3), 122-130.

Chinowsky, Paul and Molenaar, Keith (2005). “Learning organizations in construction.” *Proc., Construction Research Congress 2005: Broadening Perspectives*, ASCE, Reston, Va., 839-848.

Church, Kenneth W., and Rau, Lisa F. (1995). “Commercial applications of natural language processing.” *Commun ACM*, 38(11), 71-79.

Clough, Richard H., Sears, Clenn A., and Sears, S, Keoki. (2000). *Construction project management*, 4th Ed., Wiley, New York.

Cobb, James E., and Diekmann, James E. (1986). “Claims analysis expert system.” *Proj Manage J*, 17(2), 39-48.

Costantino, Marco, Morgan, Richard G., Collingham, Russell J., and Garigliano, Roberto (1997). “Natural language processing and information extraction: Qualitative analysis of financial news articles." *Proc., IEEE/IAFE Conference on Computational Intelligence for Financial Engineering*, IEEE, Piscataway, NJ, 116-122.

Drucker, P.R. (1993). *Post-capitalist society*, Butterworth-Heinemann, Oxford.

El-Diraby, T. E., and Kashif, K. F. (2005). "Distributed Ontology Architecture for Knowledge Management in Highway Construction." *J. Constr. Engrg. and Mgmt.*, 131(5), 591-603.

Gomez-Pereze, A. (1998). "Knowledge sharing and re-use." *Handbook of applied expert systems*, J. Liebowitz, ed., CRC, Boston.

Gómez-Pérez, Asunción, Fernández-Lopez, Mariano, and Corcho, Oscar (2004). *Ontological engineering: with examples from the areas of knowledge management, e-commerce and the semantic web*, Springer-Verlag, London.

Gruber, T.R. (1993). "A translation approach to portable ontology specification." *Knowledge Acquisition*, 5(2), 199-220.

Hachey, Ben, and Grover, Claire (2006). "Extractive summarisation of legal texts." *Artificial Intelligence and Law*, 14(4), 305-345.

Harmain, H.M., and Gaizauskas, R. (2003). "CM-Builder: A natural language-based CASE tool for object-oriented analysis." *Automated Software Engineering*, 10(2), 157-181.

Jung, Youngsoo, Kang, Seunghee, and Choi, In-Sung (2006). "Strategic knowledge management for building engineering and construction." *Proc., 2006 Architectural Engineering National Conference: Building Integration Solutions*, ASCE, Reston, Va., 56-66.

Jurafsky, Daniel, and Martin, James H. (2000). *Speech and language processing*, Prentice Hall, New Jersey.

Lame, Guiraude (2004). "Using NLP techniques to identify legal ontology components: Concepts and relations." *Artificial Intelligence and Law*, 12(4), 379-3.

Laudon, K. C., and Laudon, J. P. (1998). *Management information systems*, Prentice-Hall, Englewood Cliffs, N.J.

Lee, Tai Sik, Lee, Dong Wook, and Jee, Sang Bok (2005). "Development of Knowledge Document Management System (KDMS) for sharing construction technical documents." *Proc., Construction Research Congress 2005: Broadening Perspectives*, ASCE, Reston, Va., 1183-1191.

Lin, Yu-Cheng, Wang, Lung-Chuang, and Tserng, H. Ping (2004). "Enhancing knowledge exchange through web map-based knowledge management system in construction: Lessons learned in Taiwan." *Automation in Construction*, 15(6), 693-705.

McGee, J., and Prusak, L. (1993). *Managing information strategically*, Wiley, New York.

Nonake, Ikujiro, and Takeuchi, Hirotaka (1995). *The knowledge-creating company*, Oxford University Press, New York.

Paik, Woojin, Yilmazel, Sibel, Brown, Eric, Poulin, Maryjane, Dubon, Stephane, and Amice, Christophe (2001). "Applying natural language processing (NLP) based metadata extraction to automatically acquire user preferences." *Proc., Internat. Conf. on Knowledge Capture*, ACM, New York, NY, 116-122.

Price, David, Riloff, Ellen, Zachary, Joseph, and Harvey, Brandon (2000). “NaturalJava: A natural language interface for programming in Java.” *Proc., Int Conf Intell User Interfaces*, ACM, New York, NY, 207-211.

Raub, S., and Ruling, C. C. (2001). “The knowledge management tussle—Speech communities and rhetorical strategies in the development of knowledge management.” *J. Inf. Technol.*, 16(2), 113–130.

Rezgui, Y. (2001). “Review of information and knowledge management practices state of the art in the construction industry.” *Knowl. Eng. Rev.*, 16 2 , 241–254.

Riloff, Ellen and Phillips, William (2004). “An introduction to the Sundance and AutoSlog system.” <http://www.cs.utah.edu/~riloff/pdfs/official-sundance-tr.pdf> August 2007.

Saias, Jose, and Quaresma, Paulo (2004). “A methodology to create legal ontologies in a logic programming based web information retrieval system.” *Artificial Intelligence and Law*, 12(4), 397-417.

Senge, P.M. (1990). *The fifth Discipline: The age and practice of the learning organization*, Century Business, London.

Shamsfard, Mehrnoush, Barforoush, Ahmad Abdollahzadeh (2004). “Learning ontologies from natural language texts.” *Int. J. Human-Computer Studies*, 60(1), 17-63.

Sor, R. (2004). “Information technology and organizational structure: Vindicating theories from the past.” *Manage. Decis.*, 42(2), 316–329.

Studer, R., Benjamins, V.R., Fensel, D. (1998). “Knowledge engineering: Principles and methods.” *IEEE Transactions on Data and Knowledge Engineering*, 25(1-2), 161-197.

Toffler, A. (1990). *Powershift: Knowledge, wealth and violence at the edge of the 21st century*, Bantam Books, New York.

Turing, A.M. (1950). “Computing machinery and intelligence.” *Mind*, 59, 433-460.

Turk, Žiga (2007). “Construction informatics in European research: Topics and agendas.” *J. Comp. in Civ. Engrg.*, 21(3), 211-219.

Venters, W. (2001). “Review of the literature on knowledge management.” C-Sand working paper, <http://www.c-sand.org.uk/Documents.htm> August 2007.

Walters, Russell, Jaselskis, Edward J., and Kurtenbach, James M. (2007). “Classification of knowledge within the electrical contracting industry: A case study.” *Leadersh. Manage. Eng.*, 7(1), 11-17.

Wetherill, M., Rezgui, Y., Boddy, S., and Cooper, G. S. (2007). “Intra- and interorganizational knowledge services to promote informed sustainability practices.” *J. Comp. in Civ. Engrg.*, 21(2), 78-89.

Appendix A. Program Code

Class ConceptRelation Header File

```
// ConceptRelation class definition - ConceptRelation class public definition
// Qady - 03/04/2008

#include"sundance.h"
#include <string>
#include <vector>
using namespace std;


#ifndef CONCEPTRELATION_H
#define CONCEPTRELATION_H

class ConceptRelation
{
public:
	ConceptRelation(Sentence);
	ConceptRelation(Sentence,string);
	void setSentence(Sentence);
	void setClauseNum(string);
	string getClauseNum();
	Sentence& getSentence();
	void analyze();
	vector<vector<Constituent> >& getActiveCon();
	vector<vector<Constituent> >& getPassiveCon();
	vector<vector<Constituent> >& getRelation();
	int getCompleteRel();
	int getPartialRel();

private:
	Sentence Sent;
	string clauseNum;
	vector<vector<Constituent> > activeCon;
	vector<vector<Constituent> > passiveCon;
	vector<vector<Constituent> > relation;
	int completeRel;
	int partialRel;
};

#endif
```

Class ConceptRelation Member-function Definitions

```
#include <iostream>
#include "sundance.h"
#include <fstream>
using std::ifstream;
using std::ofstream;
#include <string>
#include <vector>
#include <iomanip>
#include "ConceptRelation.h"
using namespace std;


// ConceptRelation member-function definitions - ConceptRelation member-function implementation
// Qady - 03/04/2008

ConceptRelation::ConceptRelation(Sentence s)
{
        setSentence(s);
        setClauseNum("");
        completeRel=0;
        partialRel=0;
        analyze();
}

ConceptRelation::ConceptRelation(Sentence s, string cNum)
{
        setSentence(s);
        setClauseNum(cNum);
        completeRel=0;
        partialRel=0;
        analyze();
}

void ConceptRelation::setSentence(Sentence sent)
{
        Sent=sent;
}

Sentence& ConceptRelation::getSentence()
{
        return Sent;
}


void ConceptRelation::setClauseNum(string clNumber)
{
        clauseNum=clNumber;
}

string ConceptRelation::getClauseNum()
{
        return clauseNum;
}
```

```
void ConceptRelation::analyze()
{
	for(int i=0;i<Sent.children.length();i++){
		vector<Constituent> aCon,rel,pCon;
		int verbPos=-1;
		for(int j=0;j<Sent.children[i]->children.length();j++){//identify subject
			if(aCon.size()==0){
				if(Sent.children[i]->children[j]->role==SUBJ)
					aCon.push_back(*Sent.children[i]->children[j]);
				}
			else{//augment subject with adjacent PP
				if(Sent.children[i]->children[j]->type==2)
					aCon.push_back(*Sent.children[i]->children[j]);
				else
					break;
			}

		}

		for(int j=0;j<Sent.children[i]->children.length();j++){//identify verb
			if(Sent.children[i]->children[j]->type==1){
				rel.push_back(*Sent.children[i]->children[j]);
				verbPos=j;
				break;
			}
		}

		for(int j=0;j<Sent.children[i]->children.length();j++){//identify object
			if(pCon.size()==0){
				if(Sent.children[i]->children[j]->role==DOBJ)
					pCon.push_back(*Sent.children[i]->children[j]);
			}
			else{//augment object with adjacent PP
				if(Sent.children[i]->children[j]->type==2)
					pCon.push_back(*Sent.children[i]->children[j]);
				else
					break;
			}
		}

		if(pCon.size()==0){//if no object, look for PP after verb
			for(int j=verbPos+1;j<Sent.children[i]->children.length();j++){//if no object, look
for PP after verb
				if(pCon.size()==0){
					if(Sent.children[i]->children[j]->type==2)
						pCon.push_back(*Sent.children[i]->children[j]);
					else
						break;
				}
				else{//if object is found, augment object with adjacent PP
					if(Sent.children[i]->children[j]->type==2)
						pCon.push_back(*Sent.children[i]->children[j]);
					else
						break;
				}
```

```
				}
			}

			if(aCon.size()>0&&pCon.size()>0&&rel.size()>0)//identify number of complete and partial relations
				completeRel++;
			else
				partialRel++;

			Constituent constit;//in case any vectors are empty, add an empty Constituent to maintain vector size
			if(aCon.size()==0)
				aCon.push_back(constit);
			if(pCon.size()==0)
				pCon.push_back(constit);
			if(rel.size()==0)
				rel.push_back(constit);

			activeCon.push_back(aCon);
			passiveCon.push_back(pCon);
			relation.push_back(rel);

	}//end main for-loop
}//end member-function

vector<vector<Constituent> >& ConceptRelation::getActiveCon()
{
	return activeCon;
}

vector<vector<Constituent> >& ConceptRelation::getPassiveCon()
{
	return passiveCon;
}

vector<vector<Constituent> >& ConceptRelation::getRelation()
{
	return relation;
}

int ConceptRelation::getCompleteRel()
{
	return completeRel;
}

int ConceptRelation::getPartialRel()
{
	return partialRel;
}
```

Main Program

```
#include <iostream>
#include "sundance.h"
#include <fstream>
using std::ifstream;
using std::ofstream;
#include <string>
#include <vector>
#include <iomanip>
#include "ConceptRelation.h"
using namespace std;

int main()
{
	init_sundance();

	string clauseNum="";
	vector <ConceptRelation> conRel;
	vector <vector<Constituent> > actCon,rel,passCon;
	vector <string> clNum;

	ifstream inClientFile("test",ios::in);
	ofstream outClientFile("Analysis",ios::out);

	while(inClientFile.peek()!=EOF)
	{
		if(inClientFile.peek()==10||inClientFile.peek()==13){
			inClientFile.get();
			cout<<"enter removed"<<endl;
			}
		else if(inClientFile.peek()==60){
			getline(inClientFile,clauseNum);
			cout<<"Clause number updated: "<<clauseNum<<endl;
			}
		else{
			Sentence Sent;
			inClientFile>>Sent;
			Sent.process();
			if(Sent.children[0]->children[0]->getWords()!="Clause_End"){
				ConceptRelation con(Sent,clauseNum);
				conRel.push_back(con);
				cout<<"Concept set extracted"<<endl;
			}
			else
				cout<<"Clause End ignored"<<endl;
		}
	}

	cout<<"********************"<<endl;

	for(int i=0;i<conRel.size();i++){//build a vector of all the extracted sets
		for(int j=0;j<conRel[i].getActiveCon().size();j++){
			clNum.push_back(conRel[i].getClauseNum());
			actCon.push_back(conRel[i].getActiveCon()[j]);
			rel.push_back(conRel[i].getRelation()[j]);
```

```
                passCon.push_back(conRel[i].getPassiveCon()[j]);
            }
    }

    cout<<"Total number of concept sets: "<<clNum.size()<<endl;
    cout<<"Output generated in the following file: Analysis"<<endl;

    for(int k=0;k<clNum.size();k++){//generating output file in the required format (comma-
separated lists)
            for(int m=0;m<actCon[k].size();m++)
                outClientFile<<actCon[k][m].getWords()<<" ";
            outClientFile<<",";
            for(int m=0;m<rel[k].size();m++)
                outClientFile<<rel[k][m].getWords()<<" ";
            outClientFile<<",";
            for(int m=0;m<passCon[k].size();m++)
                outClientFile<<passCon[k][m].getWords()<<" ";
            outClientFile<<",";
            outClientFile<<clNum[k]<<endl;
            if(k!=clNum.size()-1)
                outClientFile<<",";
    }
}// end main
```

Appendix B. Input Files

Original Input File

<2.2.3>
The Owner shall furnish surveys describing physical characteristics, legal limitations and utility locations for the site of the Project, and a legal description of the site. The Contractor shall be entitled to rely on the accuracy of information furnished by the Owner but shall exercise proper precautions relating to the safe performance of the Work. Clause_End.

<4.3.1>
A Claim is a demand or assertion by one of the parties seeking, as a matter of right, adjustment or interpretation of Contract terms, payment of money, extension of time or other relief with respect to the terms of the Contract. The term "Claim" also includes other disputes and matters in question between the Owner and Contractor arising out of or relating to the Contract. Claims must be initiated by written notice. The responsibility to substantiate Claims shall rest with the party making the Claim. Clause_End.

<6.2.3>
The Owner shall be reimbursed by the Contractor for costs incurred by the Owner which are payable to a separate contractor because of delays, improperly timed activities or defective construction of the Contractor. The Owner shall be responsible to the Contractor for costs incurred by the Contractor because of delays, improperly timed activities, damage to the Work or defective construction of a separate contractor. Clause_End.

<9.8.5>
The Certificate of Substantial Completion shall be submitted to the Owner and Contractor for their written acceptance of responsibilities assigned to them in such Certificate. Upon such acceptance and consent of surety, if any, the Owner shall make payment of retainage applying to such Work or designated portion thereof. Such payment shall be adjusted for Work that is incomplete or not in accordance with the requirements of the Contract Documents. Clause_End.

<10.2.1>
The Contractor shall take reasonable precautions for safety of, and shall provide reasonable protection to prevent damage, injury or loss to:
employees on the Work and other persons who may be affected thereby;
the Work and materials and equipment to be incorporated therein, whether in storage on or off the site, under care, custody or control of the Contractor or the Contractor's Subcontractors or Sub-subcontractors; and
other property at the site or adjacent thereto, such as trees, shrubs, lawns, walks, pavements, roadways, structures and utilities not designated for removal, relocation or replacement in the course of construction. Clause_End.

<11.4.1.1>
Property insurance shall be on an "all-risk" or equivalent policy form and shall include, without limitation, insurance against the perils of fire (with extended coverage) and physical loss or damage including, without duplication of coverage, theft, vandalism, malicious mischief, collapse, earthquake, flood, windstorm, falsework, testing and startup, temporary buildings and debris removal including demolition occasioned by enforcement of any applicable legal requirements, and shall cover reasonable compensation for Architect's and Contractor's services and expenses required as a result of such insured loss. Clause_End.

Modified Input File

<2.2.3>
The Owner shall furnish surveys describing physical characteristics for the site of the Project.
The Owner shall furnish surveys describing legal limitations for the site of the Project.
The Owner shall furnish surveys describing utility locations for the site of the Project.
The Owner shall furnish a legal description of the site.
The Contractor shall be entitled to rely on the accuracy of information furnished by the Owner.
The Contractor shall exercise proper precautions relating to the safe performance of the Work.
Clause_End.

<4.3.1>
A Claim is a demand by one of the parties seeking adjustment of Contract terms.
A Claim is a demand or assertion by one of the parties seeking interpretation of Contract terms.
A Claim is a demand by one of the parties seeking payment of money.
A Claim is a demand by one of the parties seeking extension of time.
A Claim is a demand by one of the parties seeking other relief with respect to the terms of the Contract.
A Claim is an assertion by one of the parties seeking adjustment of Contract terms.
A Claim is an assertion or assertion by one of the parties seeking interpretation of Contract terms.
A Claim is an assertion by one of the parties seeking payment of money.
A Claim is an assertion by one of the parties seeking extension of time.
A Claim is an assertion by one of the parties seeking other relief with respect to the terms of the Contract.
The term "Claim" also includes other disputes in question between the Owner and Contractor.
The term "Claim" also includes other disputes arising out of the Contract.
The term "Claim" also includes other disputes relating to the Contract.
The term "Claim" also includes other matters in question between the Owner and Contractor.
The term "Claim" also includes other matters arising out of the Contract.
The term "Claim" also includes other matters relating to the Contract.
Clause_End.

<6.2.3>
The Owner shall be reimbursed by the Contractor for costs incurred by the Owner which are payable to a separate contractor because of delays.
The Owner shall be reimbursed by the Contractor for costs incurred by the Owner which are payable to a separate contractor because of improperly timed activities.
The Owner shall be reimbursed by the Contractor for costs incurred by the Owner which are payable to a separate contractor because of defective construction of the Contractor.
The Owner shall be responsible to the Contractor for costs incurred by the Contractor because of delays of a separate contractor.
The Owner shall be responsible to the Contractor for costs incurred by the Contractor because of improperly timed activities of a separate contractor.
The Owner shall be responsible to the Contractor for costs incurred by the Contractor because of damage to the Work of a separate contractor.
The Owner shall be responsible to the Contractor for costs incurred by the Contractor because of defective construction of a separate contractor.
Clause_End.

<9.8.5>
The Certificate of Substantial Completion shall be submitted to the Owner for their written acceptance of responsibilities assigned to them in such Certificate.
The Certificate of Substantial Completion shall be submitted to the Contractor for their written acceptance of responsibilities assigned to them in such Certificate.

Upon such acceptance the Owner shall make payment of retainage applying to such Work or designated portion thereof.
Upon consent of surety, the Owner shall make payment of retainage applying to such Work or designated portion thereof.
Such payment shall be adjusted for Work that is incomplete.
Such payment shall be adjusted for Work that is not in accordance with the requirements of the Contract Documents.
Clause_End.

<10.2.1>
The Contractor shall take reasonable precautions for safety of employees on the Work.
The Contractor shall take reasonable precautions for safety of other persons who may be affected thereby.
The Contractor shall take reasonable precautions for safety of the Work.
The Contractor shall take reasonable precautions for safety of the materials to be incorporated therein, in storage on the site.
The Contractor shall take reasonable precautions for safety of the materials to be incorporated therein, in storage off the site.
The Contractor shall take reasonable precautions for safety of the materials to be incorporated therein, under care of the Contractor.
The Contractor shall take reasonable precautions for safety of the materials to be incorporated therein, under care of the Contractor's Subcontractors.
The Contractor shall take reasonable precautions for safety of the materials to be incorporated therein, under care of the Contractor's Sub-subcontractors.
The Contractor shall take reasonable precautions for safety of the materials to be incorporated therein, under custody of the Contractor.
The Contractor shall take reasonable precautions for safety of the materials to be incorporated therein, under custody of the Contractor's Subcontractors.
The Contractor shall take reasonable precautions for safety of the materials to be incorporated therein, under custody of the Contractor's Sub-subcontractors.
The Contractor shall take reasonable precautions for safety of the materials to be incorporated therein, under control of the Contractor.
The Contractor shall take reasonable precautions for safety of the materials to be incorporated therein, under control of the Contractor's Subcontractors.
The Contractor shall take reasonable precautions for safety of the materials to be incorporated therein, under control of the Contractor's Sub-subcontractors.
The Contractor shall take reasonable precautions for safety of the equipment to be incorporated therein, in storage on the site.
The Contractor shall take reasonable precautions for safety of the equipment to be incorporated therein, in storage off the site.
The Contractor shall take reasonable precautions for safety of the equipment to be incorporated therein, under care of the Contractor.
The Contractor shall take reasonable precautions for safety of the equipment to be incorporated therein, under care of the Contractor's Subcontractors.
The Contractor shall take reasonable precautions for safety of the equipment to be incorporated therein, under care of the Contractor's Sub-subcontractors.
The Contractor shall take reasonable precautions for safety of the equipment to be incorporated therein, under custody of the Contractor.
The Contractor shall take reasonable precautions for safety of the equipment to be incorporated therein, under custody of the Contractor's Subcontractors.
The Contractor shall take reasonable precautions for safety of the equipment to be incorporated therein, under custody of the Contractor's Sub-subcontractors.
The Contractor shall take reasonable precautions for safety of the equipment to be incorporated therein, under control of the Contractor.

The Contractor shall take reasonable precautions for safety of the equipment to be incorporated therein, under control of the Contractor's Subcontractors.
The Contractor shall take reasonable precautions for safety of the equipment to be incorporated therein, under control of the Contractor's Sub-subcontractors.
The Contractor shall take reasonable precautions for safety of other property at the site such as trees.
The Contractor shall take reasonable precautions for safety of other property at the site such as shrubs.
The Contractor shall take reasonable precautions for safety of other property at the site such as lawns.
The Contractor shall take reasonable precautions for safety of other property at the site such as walks.
The Contractor shall take reasonable precautions for safety of other property at the site such as pavements.
The Contractor shall take reasonable precautions for safety of other property at the site such as roadways.
The Contractor shall take reasonable precautions for safety of other property at the site such as structures not designated for removal in the course of construction.
The Contractor shall take reasonable precautions for safety of other property at the site such as structures not designated for relocation in the course of construction.
The Contractor shall take reasonable precautions for safety of other property at the site such as structures not designated for replacement in the course of construction.
The Contractor shall take reasonable precautions for safety of other property at the site such as utilities not designated for removal in the course of construction.
The Contractor shall take reasonable precautions for safety of other property at the site such as utilities not designated for relocation in the course of construction.
The Contractor shall take reasonable precautions for safety of other property at the site such as utilities not designated for replacement in the course of construction.
The Contractor shall take reasonable precautions for safety of other property adjacent thereto such as trees.
The Contractor shall take reasonable precautions for safety of other property adjacent thereto such as shrubs.
The Contractor shall take reasonable precautions for safety of other property adjacent thereto such as lawns.
The Contractor shall take reasonable precautions for safety of other property adjacent thereto such as walks.
The Contractor shall take reasonable precautions for safety of other property adjacent thereto such as pavements.
The Contractor shall take reasonable precautions for safety of other property adjacent thereto such as roadways.
The Contractor shall take reasonable precautions for safety of other property adjacent thereto such as structures not designated for removal in the course of construction.
The Contractor shall take reasonable precautions for safety of other property adjacent thereto such as structures not designated for relocation in the course of construction.
The Contractor shall take reasonable precautions for safety of other property adjacent thereto such as structures not designated for replacement in the course of construction.
The Contractor shall take reasonable precautions for safety of other property adjacent thereto such as utilities not designated for removal in the course of construction.
The Contractor shall take reasonable precautions for safety of other property adjacent thereto such as utilities not designated for relocation in the course of construction.
The Contractor shall take reasonable precautions for safety of other property adjacent thereto such as utilities not designated for replacement in the course of construction.
The Contractor shall provide reasonable protection to prevent damage to employees on the Work.
The Contractor shall provide reasonable protection to prevent damage to other persons who may be affected thereby.
The Contractor shall provide reasonable protection to prevent damage to the Work.

The Contractor shall provide reasonable protection to prevent damage to the materials to be incorporated therein, in storage on the site.
The Contractor shall provide reasonable protection to prevent damage to the materials to be incorporated therein, in storage off the site.
The Contractor shall provide reasonable protection to prevent damage to the materials to be incorporated therein, under care of the Contractor.
The Contractor shall provide reasonable protection to prevent damage to the materials to be incorporated therein, under care of the Contractor's Subcontractors.
The Contractor shall provide reasonable protection to prevent damage to the materials to be incorporated therein, under care of the Contractor's Sub-subcontractors.
The Contractor shall provide reasonable protection to prevent damage to the materials to be incorporated therein, under custody of the Contractor.
The Contractor shall provide reasonable protection to prevent damage to the materials to be incorporated therein, under custody of the Contractor's Subcontractors.
The Contractor shall provide reasonable protection to prevent damage to the materials to be incorporated therein, under custody of the Contractor's Sub-subcontractors.
The Contractor shall provide reasonable protection to prevent damage to the materials to be incorporated therein, under control of the Contractor.
The Contractor shall provide reasonable protection to prevent damage to the materials to be incorporated therein, under control of the Contractor's Subcontractors.
The Contractor shall provide reasonable protection to prevent damage to the materials to be incorporated therein, under control of the Contractor's Sub-subcontractors.
The Contractor shall provide reasonable protection to prevent damage to the equipment to be incorporated therein, in storage on the site.
The Contractor shall provide reasonable protection to prevent damage to the equipment to be incorporated therein, in storage off the site.
The Contractor shall provide reasonable protection to prevent damage to the equipment to be incorporated therein, under care of the Contractor.
The Contractor shall provide reasonable protection to prevent damage to the equipment to be incorporated therein, under care of the Contractor's Subcontractors.
The Contractor shall provide reasonable protection to prevent damage to the equipment to be incorporated therein, under care of the Contractor's Sub-subcontractors.
The Contractor shall provide reasonable protection to prevent damage to the equipment to be incorporated therein, under custody of the Contractor.
The Contractor shall provide reasonable protection to prevent damage to the equipment to be incorporated therein, under custody of the Contractor's Subcontractors.
The Contractor shall provide reasonable protection to prevent damage to the equipment to be incorporated therein, under custody of the Contractor's Sub-subcontractors.
The Contractor shall provide reasonable protection to prevent damage to the equipment to be incorporated therein, under control of the Contractor.
The Contractor shall provide reasonable protection to prevent damage to the equipment to be incorporated therein, under control of the Contractor's Subcontractors.
The Contractor shall provide reasonable protection to prevent damage to the equipment to be incorporated therein, under control of the Contractor's Sub-subcontractors.
The Contractor shall provide reasonable protection to prevent damage to other property at the site such as trees.
The Contractor shall provide reasonable protection to prevent damage to other property at the site such as shrubs.
The Contractor shall provide reasonable protection to prevent damage to other property at the site such as lawns.
The Contractor shall provide reasonable protection to prevent damage to other property at the site such as walks.
The Contractor shall provide reasonable protection to prevent damage to other property at the site such as pavements.

The Contractor shall provide reasonable protection to prevent damage to other property at the site such as roadways.
The Contractor shall provide reasonable protection to prevent damage to other property at the site such as structures not designated for removal in the course of construction.
The Contractor shall provide reasonable protection to prevent damage to other property at the site such as structures not designated for relocation in the course of construction.
The Contractor shall provide reasonable protection to prevent damage to other property at the site such as structures not designated for replacement in the course of construction.
The Contractor shall provide reasonable protection to prevent damage to other property at the site such as utilities not designated for removal in the course of construction.
The Contractor shall provide reasonable protection to prevent damage to other property at the site such as utilities not designated for relocation in the course of construction.
The Contractor shall provide reasonable protection to prevent damage to other property at the site such as utilities not designated for replacement in the course of construction.
The Contractor shall provide reasonable protection to prevent damage to other property adjacent thereto such as trees.
The Contractor shall provide reasonable protection to prevent damage to other property adjacent thereto such as shrubs.
The Contractor shall provide reasonable protection to prevent damage to other property adjacent thereto such as lawns.
The Contractor shall provide reasonable protection to prevent damage to other property adjacent thereto such as walks.
The Contractor shall provide reasonable protection to prevent damage to other property adjacent thereto such as pavements.
The Contractor shall provide reasonable protection to prevent damage to other property adjacent thereto such as roadways.
The Contractor shall provide reasonable protection to prevent damage to other property adjacent thereto such as structures not designated for removal in the course of construction.
The Contractor shall provide reasonable protection to prevent damage to other property adjacent thereto such as structures not designated for relocation in the course of construction.
The Contractor shall provide reasonable protection to prevent damage to other property adjacent thereto such as structures not designated for replacement in the course of construction.
The Contractor shall provide reasonable protection to prevent damage to other property adjacent thereto such as utilities not designated for removal in the course of construction.
The Contractor shall provide reasonable protection to prevent damage to other property adjacent thereto such as utilities not designated for relocation in the course of construction.
The Contractor shall provide reasonable protection to prevent damage to other property adjacent thereto such as utilities not designated for replacement in the course of construction.
The Contractor shall provide reasonable protection to prevent injury to employees on the Work.
The Contractor shall provide reasonable protection to prevent injury to other persons who may be affected thereby.
The Contractor shall provide reasonable protection to prevent injury to the Work.
The Contractor shall provide reasonable protection to prevent injury to the materials to be incorporated therein, in storage on the site.
The Contractor shall provide reasonable protection to prevent injury to the materials to be incorporated therein, in storage off the site.
The Contractor shall provide reasonable protection to prevent injury to the materials to be incorporated therein, under care of the Contractor.
The Contractor shall provide reasonable protection to prevent injury to the materials to be incorporated therein, under care of the Contractor's Subcontractors.
The Contractor shall provide reasonable protection to prevent injury to the materials to be incorporated therein, under care of the Contractor's Sub-subcontractors.
The Contractor shall provide reasonable protection to prevent injury to the materials to be incorporated therein, under custody of the Contractor.

The Contractor shall provide reasonable protection to prevent injury to the materials to be incorporated therein, under custody of the Contractor's Subcontractors.
The Contractor shall provide reasonable protection to prevent injury to the materials to be incorporated therein, under custody of the Contractor's Sub-subcontractors.
The Contractor shall provide reasonable protection to prevent injury to the materials to be incorporated therein, under control of the Contractor.
The Contractor shall provide reasonable protection to prevent injury to the materials to be incorporated therein, under control of the Contractor's Subcontractors.
The Contractor shall provide reasonable protection to prevent injury to the materials to be incorporated therein, under control of the Contractor's Sub-subcontractors.
The Contractor shall provide reasonable protection to prevent injury to the equipment to be incorporated therein, in storage on the site.
The Contractor shall provide reasonable protection to prevent injury to the equipment to be incorporated therein, in storage off the site.
The Contractor shall provide reasonable protection to prevent injury to the equipment to be incorporated therein, under care of the Contractor.
The Contractor shall provide reasonable protection to prevent injury to the equipment to be incorporated therein, under care of the Contractor's Subcontractors.
The Contractor shall provide reasonable protection to prevent injury to the equipment to be incorporated therein, under care of the Contractor's Sub-subcontractors.
The Contractor shall provide reasonable protection to prevent injury to the equipment to be incorporated therein, under custody of the Contractor.
The Contractor shall provide reasonable protection to prevent injury to the equipment to be incorporated therein, under custody of the Contractor's Subcontractors.
The Contractor shall provide reasonable protection to prevent injury to the equipment to be incorporated therein, under custody of the Contractor's Sub-subcontractors.
The Contractor shall provide reasonable protection to prevent injury to the equipment to be incorporated therein, under control of the Contractor.
The Contractor shall provide reasonable protection to prevent injury to the equipment to be incorporated therein, under control of the Contractor's Subcontractors.
The Contractor shall provide reasonable protection to prevent injury to the equipment to be incorporated therein, under control of the Contractor's Sub-subcontractors.
The Contractor shall provide reasonable protection to prevent injury to other property at the site such as trees.
The Contractor shall provide reasonable protection to prevent injury to other property at the site such as shrubs.
The Contractor shall provide reasonable protection to prevent injury to other property at the site such as lawns.
The Contractor shall provide reasonable protection to prevent injury to other property at the site such as walks.
The Contractor shall provide reasonable protection to prevent injury to other property at the site such as pavements.
The Contractor shall provide reasonable protection to prevent injury to other property at the site such as roadways.
The Contractor shall provide reasonable protection to prevent injury to other property at the site such as structures not designated for removal in the course of construction.
The Contractor shall provide reasonable protection to prevent injury to other property at the site such as structures not designated for relocation in the course of construction.
The Contractor shall provide reasonable protection to prevent injury to other property at the site such as structures not designated for replacement in the course of construction.
The Contractor shall provide reasonable protection to prevent injury to other property at the site such as utilities not designated for removal in the course of construction.
The Contractor shall provide reasonable protection to prevent injury to other property at the site such as utilities not designated for relocation in the course of construction.

The Contractor shall provide reasonable protection to prevent injury to other property at the site such as utilities not designated for replacement in the course of construction.
The Contractor shall provide reasonable protection to prevent injury to other property adjacent thereto such as trees.
The Contractor shall provide reasonable protection to prevent injury to other property adjacent thereto such as shrubs.
The Contractor shall provide reasonable protection to prevent injury to other property adjacent thereto such as lawns.
The Contractor shall provide reasonable protection to prevent injury to other property adjacent thereto such as walks.
The Contractor shall provide reasonable protection to prevent injury to other property adjacent thereto such as pavements.
The Contractor shall provide reasonable protection to prevent injury to other property adjacent thereto such as roadways.
The Contractor shall provide reasonable protection to prevent injury to other property adjacent thereto such as structures not designated for removal in the course of construction.
The Contractor shall provide reasonable protection to prevent injury to other property adjacent thereto such as structures not designated for relocation in the course of construction.
The Contractor shall provide reasonable protection to prevent injury to other property adjacent thereto such as structures not designated for replacement in the course of construction.
The Contractor shall provide reasonable protection to prevent injury to other property adjacent thereto such as utilities not designated for removal in the course of construction.
The Contractor shall provide reasonable protection to prevent injury to other property adjacent thereto such as utilities not designated for relocation in the course of construction.
The Contractor shall provide reasonable protection to prevent injury to other property adjacent thereto such as utilities not designated for replacement in the course of construction.
The Contractor shall provide reasonable protection to prevent loss to employees on the Work.
The Contractor shall provide reasonable protection to prevent loss to other persons who may be affected thereby.
The Contractor shall provide reasonable protection to prevent loss to the Work.
The Contractor shall provide reasonable protection to prevent loss to the materials to be incorporated therein, in storage on the site.
The Contractor shall provide reasonable protection to prevent loss to the materials to be incorporated therein, in storage off the site.
The Contractor shall provide reasonable protection to prevent loss to the materials to be incorporated therein, under care of the Contractor.
The Contractor shall provide reasonable protection to prevent loss to the materials to be incorporated therein, under care of the Contractor's Subcontractors.
The Contractor shall provide reasonable protection to prevent loss to the materials to be incorporated therein, under care of the Contractor's Sub-subcontractors.
The Contractor shall provide reasonable protection to prevent loss to the materials to be incorporated therein, under custody of the Contractor.
The Contractor shall provide reasonable protection to prevent loss to the materials to be incorporated therein, under custody of the Contractor's Subcontractors.
The Contractor shall provide reasonable protection to prevent loss to the materials to be incorporated therein, under custody of the Contractor's Sub-subcontractors.
The Contractor shall provide reasonable protection to prevent loss to the materials to be incorporated therein, under control of the Contractor.
The Contractor shall provide reasonable protection to prevent loss to the materials to be incorporated therein, under control of the Contractor's Subcontractors.
The Contractor shall provide reasonable protection to prevent loss to the materials to be incorporated therein, under control of the Contractor's Sub-subcontractors.
The Contractor shall provide reasonable protection to prevent loss to the equipment to be incorporated therein, in storage on the site.

The Contractor shall provide reasonable protection to prevent loss to the equipment to be incorporated therein, in storage off the site.
The Contractor shall provide reasonable protection to prevent loss to the equipment to be incorporated therein, under care of the Contractor.
The Contractor shall provide reasonable protection to prevent loss to the equipment to be incorporated therein, under care of the Contractor's Subcontractors.
The Contractor shall provide reasonable protection to prevent loss to the equipment to be incorporated therein, under care of the Contractor's Sub-subcontractors.
The Contractor shall provide reasonable protection to prevent loss to the equipment to be incorporated therein, under custody of the Contractor.
The Contractor shall provide reasonable protection to prevent loss to the equipment to be incorporated therein, under custody of the Contractor's Subcontractors.
The Contractor shall provide reasonable protection to prevent loss to the equipment to be incorporated therein, under custody of the Contractor's Sub-subcontractors.
The Contractor shall provide reasonable protection to prevent loss to the equipment to be incorporated therein, under control of the Contractor.
The Contractor shall provide reasonable protection to prevent loss to the equipment to be incorporated therein, under control of the Contractor's Subcontractors.
The Contractor shall provide reasonable protection to prevent loss to the equipment to be incorporated therein, under control of the Contractor's Sub-subcontractors.
The Contractor shall provide reasonable protection to prevent loss to other property at the site such as trees.
The Contractor shall provide reasonable protection to prevent loss to other property at the site such as shrubs.
The Contractor shall provide reasonable protection to prevent loss to other property at the site such as lawns.
The Contractor shall provide reasonable protection to prevent loss to other property at the site such as walks.
The Contractor shall provide reasonable protection to prevent loss to other property at the site such as pavements.
The Contractor shall provide reasonable protection to prevent loss to other property at the site such as roadways.
The Contractor shall provide reasonable protection to prevent loss to other property at the site such as structures not designated for removal in the course of construction.
The Contractor shall provide reasonable protection to prevent loss to other property at the site such as structures not designated for relocation in the course of construction.
The Contractor shall provide reasonable protection to prevent loss to other property at the site such as structures not designated for replacement in the course of construction.
The Contractor shall provide reasonable protection to prevent loss to other property at the site such as utilities not designated for removal in the course of construction.
The Contractor shall provide reasonable protection to prevent loss to other property at the site such as utilities not designated for relocation in the course of construction.
The Contractor shall provide reasonable protection to prevent loss to other property at the site such as utilities not designated for replacement in the course of construction.
The Contractor shall provide reasonable protection to prevent loss to other property adjacent thereto such as trees.
The Contractor shall provide reasonable protection to prevent loss to other property adjacent thereto such as shrubs.
The Contractor shall provide reasonable protection to prevent loss to other property adjacent thereto such as lawns.
The Contractor shall provide reasonable protection to prevent injury to other property adjacent thereto such as walks.
The Contractor shall provide reasonable protection to prevent loss to other property adjacent thereto such as pavements.

The Contractor shall provide reasonable protection to prevent loss to other property adjacent thereto such as roadways.
The Contractor shall provide reasonable protection to prevent loss to other property adjacent thereto such as structures not designated for removal in the course of construction.
The Contractor shall provide reasonable protection to prevent loss to other property adjacent thereto such as structures not designated for relocation in the course of construction.
The Contractor shall provide reasonable protection to prevent loss to other property adjacent thereto such as structures not designated for replacement in the course of construction.
The Contractor shall provide reasonable protection to prevent loss to other property adjacent thereto such as utilities not designated for removal in the course of construction.
The Contractor shall provide reasonable protection to prevent loss to other property adjacent thereto such as utilities not designated for relocation in the course of construction.
The Contractor shall provide reasonable protection to prevent loss to other property adjacent thereto such as utilities not designated for replacement in the course of construction.
Clause_End.

<11.4.1.1>
Property insurance shall be on an "all-risk" or equivalent policy form.
Property insurance shall include insurance against the perils of fire (with extended coverage).
Property insurance shall include physical loss including theft.
Property insurance shall include physical loss including vandalism.
Property insurance shall include physical loss including malicious mischief.
Property insurance shall include physical loss including collapse.
Property insurance shall include physical loss including earthquake.
Property insurance shall include physical loss including flood.
Property insurance shall include physical loss including windstorm.
Property insurance shall include physical loss including falsework.
Property insurance shall include physical loss including testing and startup.
Property insurance shall include physical loss including temporary buildings.
Property insurance shall include physical loss including debris removal including demolition occasioned by enforcement of any applicable legal requirements.
Property insurance shall include physical damage including theft.
Property insurance shall include physical damage including vandalism.
Property insurance shall include physical damage including malicious mischief.
Property insurance shall include physical damage including collapse.
Property insurance shall include physical damage including earthquake.
Property insurance shall include physical damage including flood.
Property insurance shall include physical damage including windstorm.
Property insurance shall include physical damage including falsework.
Property insurance shall include physical damage including testing and startup.
Property insurance shall include physical damage including temporary buildings.
Property insurance shall include physical damage including debris removal including demolition occasioned by enforcement of any applicable legal requirements.
Property insurance shall cover reasonable compensation for Architect's services required as a result of such insured loss.
Property insurance shall cover reasonable compensation for Architect's expenses required as a result of such insured loss.
Property insurance shall cover reasonable compensation for Contractor's services required as a result of such insured loss.
Property insurance shall cover reasonable compensation for Contractor's expenses required as a result of such insured loss.
Clause_End.

Appendix C. System Output

Output1

Ser.	Section #	Active Concept	Relation	Passive Concept
1	<2.2.3>	The Owner	shall furnish	surveys
2	<2.2.3>	surveys	describing	physical characteristics
3	<2.2.3>	The Contractor	shall be entitled to rely	on the accuracy of information
4	<2.2.3>	the accuracy of information	furnished	by the Owner
5	<2.2.3>	the Owner	shall exercise	Proper precautions relating to the safe performance of the Work
6	<4.3.1>	A Claim	is	a demand or assertion by one of the parties seeking
7	<4.3.1>	a matter of right	Contract	terms
8	<4.3.1>	The term Claim	also includes	other
9	<4.3.1>	other	disputes	
10	<4.3.1>	other	matters	in question between the Owner and Contractor arising
11	<4.3.1>	Claims	must be initiated	by written notice
12	<4.3.1>	The responsibility	to substantiate	Claims
13	<4.3.1>	the party	making	the Claim
14	<6.2.3>	The Owner	shall be reimbursed	by the Contractor for costs
15	<6.2.3>	costs	incurred	by the Owner
16	<6.2.3>	the Owner	are	
17	<6.2.3>	the Owner	improperly timed	activities or defective construction of the Contractor
18	<6.2.3>	The Owner	shall be	
19	<6.2.3>	costs	incurred	by the Contractor BECAUSE_OF delays
20	<6.2.3>	costs	improperly timed	activities
21	<6.2.3>	costs	damage	to the Work or defective construction of a separate contractor
22	<9.8.5>	The Certificate of Substantial Completion	shall be submitted	to the Owner and Contractor for their written acceptance of responsibilities
23	<9.8.5>	their written acceptance of responsibilities	assigned	to them in such Certificate

Table: Output1 (continued)

Ser.	Section #	Active Concept	Relation	Passive Concept
24	<9.8.5>	such acceptance	consent	of surety
25	<9.8.5>	any	shall make	payment of retainage applying to such Work
26	<9.8.5>	any	designated	portion thereof
27	<9.8.5>	Such payment	shall be adjusted	for Work
28	<9.8.5>	Work	is	
29	<10.2.1>	The Contractor	shall take	reasonable precautions for safety
30	<10.2.1>	The Contractor	shall provide	reasonable protection
31	<10.2.1>	the Work and other persons	may be affected thereby	
32	<10.2.1>	the Work and materials and equipment	to be incorporated	Therein
33	<10.2.1>	custody or control of the Contractor or the @Contractor@s Subcontractors or Sub-subcontractors		
34	<10.2.1>	other property at the site or adjacent thereto	walks	
35	<10.2.1>	other property	structures	
36	<10.2.1>	utilities	not designated	for removal
37	<11.4.1.1>	Property insurance	shall be	on an all-risk or equivalent policy form
38	<11.4.1.1>	Property insurance	shall include	
39	<11.4.1.1>	Property insurance	damage including	
40	<11.4.1.1>	Property insurance	collapse	
41	<11.4.1.1>	debris removal	including	demolition
42	<11.4.1.1>	demolition	occasioned	by enforcement of any applicable legal requirements
43	<11.4.1.1>	demolition	shall cover	reasonable compensation for @Architect@s and @Contractor@s services and expenses
44	<11.4.1.1>	expenses	required	as a result of such insured loss

Output2

Ser.	Section #	Active Concept	Relation	Passive Concept
1	<2.2.3>	The Owner	shall furnish	surveys
2	<2.2.3>	surveys	describing	physical characteristics for the site of the Project
3	<2.2.3>	surveys	describing	legal limitations for the site of the Project
4	<2.2.3>	The Owner	shall furnish	surveys describing utility locations for the site of the Project
5	<2.2.3>	The Owner	shall furnish	a legal description of the site
6	<2.2.3>	The Contractor	shall be entitled to rely	on the accuracy of information
7	<2.2.3>	the accuracy of information	furnished	by the Owner
8	<2.2.3>	The Contractor	shall exercise	proper precautions relating to the safe performance of the Work
9	<4.3.1>	A Claim	is	a demand by one of the parties seeking adjustment
10	<4.3.1>	one of the parties seeking adjustment	Contract	terms
11	<4.3.1>	A Claim	is	a demand by one of the parties seeking interpretation
12	<4.3.1>	one of the parties seeking interpretation	Contract	terms
13	<4.3.1>	A Claim	is	a demand by one of the parties seeking payment of money
14	<4.3.1>	A Claim	is	a demand by one of the parties seeking extension of time
15	<4.3.1>	A Claim	is	a demand by one of the parties seeking other relief with respect to the terms of the Contract
16	<4.3.1>	A Claim	is	an assertion by one of the parties seeking adjustment

Table: Output2 (continued)

Ser.	Section #	Active Concept	Relation	Passive Concept
17	<4.3.1>	A Claim	is	an assertion or assertion by one of the parties seeking interpretation
18	<4.3.1>	A Claim	is	an assertion by one of the parties seeking payment of money
19	<4.3.1>	A Claim	is	an assertion by one of the parties seeking extension of time
20	<4.3.1>	A Claim	is	an assertion by one of the parties seeking other relief with respect to the terms of the Contract
21	<4.3.1>	The term Claim	also includes	other
22	<4.3.1>	other	disputes	in question between the Owner and Contractor
23	<4.3.1>	other	disputes	arising
24	<4.3.1>	other	disputes	relating to the Contract
25	<4.3.1>	other	matters	in question between the Owner and Contractor
26	<4.3.1>	other	matters	arising
27	<4.3.1>	other	matters	relating to the Contract
28	<4.3.1>	Claims	must be initiated	by written notice
29	<4.3.1>	The responsibility	to substantiate	Claims
30	<4.3.1>	the party	making	the Claim
31	<6.2.3>	The Owner	shall be reimbursed	by the Contractor for costs
32	<6.2.3>	costs	incurred	by the Owner
33	<6.2.3>	the Owner	are	
34	<6.2.3>	The Owner	shall be	
35	<6.2.3>	costs	incurred	by the Contractor BECAUSE_OF delays of a separate contractor
36	<6.2.3>	costs	incurred	by the Contractor BECAUSE_OF improperly timed activities of a separate contractor

Table: Output2 (continued)

Ser.	Section #	Active Concept	Relation	Passive Concept
37	<6.2.3>	costs	incurred	by the Contractor BECAUSE_OF damage to the Work of a separate contractor
38	<6.2.3>	costs	incurred	by the Contractor BECAUSE_OF defective construction of a separate contractor
39	<9.8.5>	The Certificate of Substantial Completion	shall be submitted	to the Owner for their written acceptance of responsibilities
40	<9.8.5>	their written acceptance of responsibilities	assigned	to them in such Certificate
41	<9.8.5>	The Certificate of Substantial Completion	shall be submitted	to the Contractor for their written acceptance of responsibilities
42	<9.8.5>	the Owner	shall make	payment of retainage applying to such Work
43	<9.8.5>	the Owner	designated	portion thereof
44	<9.8.5>		consent	of surety
45	<9.8.5>	Such payment	shall be adjusted	for Work
46	<9.8.5>	Work	is	
47	<9.8.5>	Work	is not	in accordance with the requirements of the Contract Documents
48	<10.2.1>	The Contractor	shall take	reasonable precautions for safety of employees on the Work
49	<10.2.1>	The Contractor	shall take	reasonable precautions for safety of other persons
50	<10.2.1>	safety of other persons	may be affected thereby	
51	<10.2.1>	The Contractor	shall take	reasonable precautions for safety of the Work
52	<10.2.1>	The Contractor	shall take	reasonable precautions for safety of the materials
53	<10.2.1>	safety of the materials	to be incorporated	therein
54	<10.2.1>	The Contractor	shall take	reasonable precautions for safety of the equipment

Table: Output2 (continued)

Ser.	Section #	Active Concept	Relation	Passive Concept
55	<10.2.1>	safety of the equipment	to be incorporated	therein
56	<10.2.1>	The Contractor	shall take	reasonable precautions for safety of other property at the site SUCH_AS trees
57	<10.2.1>	The Contractor	shall take	reasonable precautions for safety of other property at the site SUCH_AS shrubs
58	<10.2.1>	The Contractor	shall take	reasonable precautions for safety of other property at the site SUCH_AS lawns
59	<10.2.1>	The Contractor	shall take	reasonable precautions for safety of other property at the site
60	<10.2.1>	the site	walks	
61	<10.2.1>	The Contractor	shall take	reasonable precautions for safety of other property at the site SUCH_AS pavements
62	<10.2.1>	The Contractor	shall take	reasonable precautions for safety of other property at the site SUCH_AS roadways
63	<10.2.1>	The Contractor	shall take	reasonable precautions for safety of other property at the site SUCH_AS structures
64	<10.2.1>	structures	not designated	for removal in the course of construction
65	<10.2.1>	structures	not designated	for relocation in the course of construction
66	<10.2.1>	structures	not designated	for replacement in the course of construction
67	<10.2.1>	The Contractor	shall take	reasonable precautions for safety of other property at the site SUCH_AS utilities
68	<10.2.1>	utilities	not designated	for removal in the course of construction
69	<10.2.1>	utilities	not designated	for relocation in the course of construction

Table: Output2 (continued)

Ser.	Section #	Active Concept	Relation	Passive Concept
70	<10.2.1>	utilities	not designated	for replacement in the course of construction
71	<10.2.1>	The Contractor	shall take	reasonable precautions for safety of other property adjacent thereto SUCH_AS trees
72	<10.2.1>	The Contractor	shall take	reasonable precautions for safety of other property adjacent thereto SUCH_AS shrubs
73	<10.2.1>	The Contractor	shall take	reasonable precautions for safety of other property adjacent thereto SUCH_AS lawns
74	<10.2.1>	The Contractor	shall take	reasonable precautions for safety of other property adjacent thereto
75	<10.2.1>	safety of other property adjacent thereto	walks	
76	<10.2.1>	The Contractor	shall take	reasonable precautions for safety of other property adjacent thereto SUCH_AS pavements
77	<10.2.1>	The Contractor	shall take	reasonable precautions for safety of other property adjacent thereto SUCH_AS roadways
78	<10.2.1>	The Contractor	shall take	reasonable precautions for safety of other property adjacent thereto SUCH_AS structures
79	<10.2.1>	The Contractor	shall take	reasonable precautions for safety of other property adjacent thereto SUCH_AS utilities

Table: Output2 (continued)

Ser.	Section #	Active Concept	Relation	Passive Concept
80	<10.2.1>	The Contractor	shall provide	reasonable protection
81	<10.2.1>	other persons	may be affected thereby	
82	<10.2.1>	the materials	to be incorporated	therein
83	<10.2.1>	the equipment	to be incorporated	therein
84	<10.2.1>	other property adjacent thereto	walks	
85	<11.4.1.1>	Property insurance	shall be	on an all-risk or equivalent policy form
86	<11.4.1.1>	Property insurance	shall include	insurance against the perils of fire
87	<11.4.1.1>	Property insurance	shall include	physical loss
88	<11.4.1.1>	physical loss	including	theft
89	<11.4.1.1>	physical loss	collapse	
90	<11.4.1.1>	debris removal	including	demolition
91	<11.4.1.1>	demolition	occasioned	by enforcement of any applicable legal requirements
92	<11.4.1.1>	physical loss	including	vandalism
93	<11.4.1.1>	physical loss	including	malicious mischief
94	<11.4.1.1>	physical loss	including collapse	
95	<11.4.1.1>	physical loss	including	earthquake
96	<11.4.1.1>	physical loss	including	flood
97	<11.4.1.1>	physical loss	including	windstorm
98	<11.4.1.1>	physical loss	including	falsework
99	<11.4.1.1>	physical loss	including	testing and startup
100	<11.4.1.1>	physical loss	including	temporary buildings
101	<11.4.1.1>	physical loss	including	debris removal
102	<11.4.1.1>	Property insurance	shall include	physical damage
103	<11.4.1.1>	physical damage	including	theft
104	<11.4.1.1>	physical damage	including	vandalism
105	<11.4.1.1>	physical damage	including	malicious mischief
106	<11.4.1.1>	physical damage	including collapse	
107	<11.4.1.1>	physical damage	including	earthquake
108	<11.4.1.1>	physical damage	including	flood
109	<11.4.1.1>	physical damage	including	windstorm
110	<11.4.1.1>	physical damage	including	falsework
111	<11.4.1.1>	physical damage	including	testing and startup
112	<11.4.1.1>	physical damage	including	temporary buildings
113	<11.4.1.1>	physical damage	including	debris removal

Table: Output2 (continued)

Ser.	Section #	Active Concept	Relation	Passive Concept
114	<11.4.1.1>	Property insurance	shall cover	reasonable compensation for @Architect@s services
115	<11.4.1.1>	@Architect@s services	required	as a result of such insured loss
116	<11.4.1.1>	Property insurance	shall cover	reasonable compensation for @Architect@s expenses
117	<11.4.1.1>	@Architect@s expenses	required	as a result of such insured loss
118	<11.4.1.1>	Property insurance	shall cover	reasonable compensation for @Contractor@s services
119	<11.4.1.1>	@Contractor@s services	required	as a result of such insured loss
120	<11.4.1.1>	Property insurance	shall cover	reasonable compensation for @Contractor@s expenses
121	<11.4.1.1>	@Contractor@s expenses	required	as a result of such insured loss

Appendix D. Evaluation Exercise

1. The Owner shall furnish surveys describing physical characteristics, legal limitations and utility locations for the site of the Project, and a legal description of the site. The Contractor shall be entitled to rely on the accuracy of information furnished by the Owner but shall exercise proper precautions relating to the safe performance of the Work.

Active Concept	Relation	Passive Concept

2. The Owner shall be reimbursed by the Contractor for costs incurred by the Owner which are payable to a separate contractor because of delays, improperly timed activities or defective construction of the Contractor. The Owner shall be responsible to the Contractor for costs incurred by the Contractor because of delays, improperly timed activities, damage to the Work or defective construction of a separate contractor.

Active Concept	Relation	Passive Concept

3. The Certificate of Substantial Completion shall be submitted to the Owner and Contractor for their written acceptance of responsibilities assigned to them in such Certificate. Upon such acceptance and consent of surety, if any, the Owner shall make payment of retainage applying to such Work or designated portion thereof. Such payment shall be adjusted for Work that is incomplete or not in accordance with the requirements of the Contract Documents.

Active Concept	Relation	Passive Concept

4. Property insurance shall be on an "all-risk" or equivalent policy form and shall include, without limitation, insurance against the perils of fire (with extended coverage) and physical loss or damage including, without duplication of coverage, theft, vandalism, malicious mischief, collapse, earthquake, flood, windstorm, falsework, testing and startup, temporary buildings and debris removal including demolition occasioned by enforcement of any applicable legal requirements, and shall cover reasonable compensation for Architect's and Contractor's services and expenses required as a result of such insured loss.

Active Concept	Relation	Passive Concept

5. A Claim is a demand or assertion by one of the parties seeking, as a matter of right, adjustment or interpretation of Contract terms, payment of money, extension of time or other relief with respect to the terms of the Contract. The term "Claim" also includes other disputes and matters in question between the Owner and Contractor arising out of or relating to the Contract. Claims must be initiated by written notice. The responsibility to substantiate Claims shall rest with the party making the Claim.

Active Concept	Relation	Passive Concept

6. The Contractor shall take reasonable precautions for safety of, and shall provide reasonable protection to prevent damage, injury or loss to:
   employees on the Work and other persons who may be affected thereby;
   the Work and materials and equipment to be incorporated therein, whether in storage on or off the site, under care, custody or control of the Contractor or the Contractor's Subcontractors or Sub-subcontractors; and
   other property at the site or adjacent thereto, such as trees, shrubs, lawns, walks, pavements, roadways, structures and utilities not designated for removal, relocation or replacement in the course of construction.

Active Concept	Relation	Passive Concept

Appendix E. Kappa Score Results

Comparison	Section #	K	Average K
Evaluator A vs Evaluator B	2.2.3	0.329	0.280
	6.2.3	0.233	
	9.8.5	0.426	
	11.4.1.1	0.375	
	4.3.1	0.295	
	10.2.1	0.021	
Evaluator A vs Evaluator C	2.2.3	0.358	0.403
	6.2.3	0.394	
	9.8.5	0.394	
	11.4.1.1	0.500	
	4.3.1	0.749	
	10.2.1	0.021	
Evaluator A vs Evaluator D	2.2.3	0.438	0.394
	6.2.3	0.329	
	9.8.5	0.492	
	11.4.1.1	0.232	
	4.3.1	0.845	
	10.2.1	0.030	
Evaluator A vs Evaluator E	2.2.3	0.267	0.353
	6.2.3	0.580	
	9.8.5	0.550	
	11.4.1.1	0.194	
	4.3.1	0.475	
	10.2.1	0.053	
Evaluator A vs Evaluator F	2.2.3	0.149	0.289
	6.2.3	0.394	
	9.8.5	0.495	
	11.4.1.1	0.146	
	4.3.1	0.528	
	10.2.1	0.022	
Evaluator A vs Evaluator G	2.2.3	0.394	0.433
	6.2.3	0.394	
	9.8.5	0.596	
	11.4.1.1	0.575	
	4.3.1	0.398	
	10.2.1	0.242	
Evaluator A vs Output1	2.2.3	0.213	0.098
	6.2.3	0.059	
	9.8.5	0.161	
	11.4.1.1	0.026	
	4.3.1	0.103	
	10.2.1	0.028	

Table: Kappa score results (continued)

Comparison	Section #	K	Average K
Evaluator A vs Output2	2.2.3	0.550	0.343
	6.2.3	0.304	
	9.8.5	0.329	
	11.4.1.1	0.675	
	4.3.1	0.093	
	10.2.1	0.110	
Evaluator B vs Evaluator C	2.2.3	0.899	0.469
	6.2.3	0.413	
	9.8.5	0.536	
	11.4.1.1	0.533	
	4.3.1	0.321	
	10.2.1	0.113	
Evaluator B vs Evaluator D	2.2.3	0.697	0.289
	6.2.3	0.184	
	9.8.5	0.282	
	11.4.1.1	0.142	
	4.3.1	0.319	
	10.2.1	0.110	
Evaluator B vs Evaluator E	2.2.3	0.633	0.282
	6.2.3	0.292	
	9.8.5	0.330	
	11.4.1.1	0.151	
	4.3.1	0.176	
	10.2.1	0.111	
Evaluator B vs Evaluator F	2.2.3	0.580	0.302
	6.2.3	0.413	
	9.8.5	0.233	
	11.4.1.1	0.171	
	4.3.1	0.249	
	10.2.1	0.166	
Evaluator B vs Evaluator G	2.2.3	0.899	0.482
	6.2.3	0.413	
	9.8.5	0.690	
	11.4.1.1	0.566	
	4.3.1	0.232	
	10.2.1	0.093	

Table: Kappa score results (continued)

Comparison	Section #	K	Average K
Evaluator B vs Output1	2.2.3	0.495	0.203
	6.2.3	0.174	
	9.8.5	0.174	
	11.4.1.1	0.137	
	4.3.1	0.110	
	10.2.1	0.131	
Evaluator B vs Output2	2.2.3	0.633	0.307
	6.2.3	0.174	
	9.8.5	0.497	
	11.4.1.1	0.348	
	4.3.1	0.099	
	10.2.1	0.091	
Evaluator C vs Evaluator D	2.2.3	0.775	0.485
	6.2.3	0.492	
	9.8.5	0.492	
	11.4.1.1	0.271	
	4.3.1	0.732	
	10.2.1	0.146	
Evaluator C vs Evaluator E	2.2.3	0.697	0.382
	6.2.3	0.495	
	9.8.5	0.550	
	11.4.1.1	0.213	
	4.3.1	0.221	
	10.2.1	0.113	
Evaluator C vs Evaluator F	2.2.3	0.633	0.479
	6.2.3	1.000	
	9.8.5	0.358	
	11.4.1.1	0.110	
	4.3.1	0.587	
	10.2.1	0.188	
Evaluator C vs Evaluator G	2.2.3	1.000	0.759
	6.2.3	1.000	
	9.8.5	0.775	
	11.4.1.1	0.826	
	4.3.1	0.764	
	10.2.1	0.192	

Table: Kappa score results (continued)

Comparison	Section #	K	Average K
Evaluator C vs Output1	2.2.3	0.394	0.230
	6.2.3	0.293	
	9.8.5	0.267	
	11.4.1.1	0.114	
	4.3.1	0.208	
	10.2.1	0.105	
Evaluator C vs Output2	2.2.3	0.697	0.401
	6.2.3	0.293	
	9.8.5	0.596	
	11.4.1.1	0.500	
	4.3.1	0.155	
	10.2.1	0.164	
Evaluator D vs Evaluator E	2.2.3	0.495	0.366
	6.2.3	0.304	
	9.8.5	0.325	
	11.4.1.1	0.536	
	4.3.1	0.207	
	10.2.1	0.330	
Evaluator D vs Evaluator F	2.2.3	0.450	0.370
	6.2.3	0.492	
	9.8.5	0.175	
	11.4.1.1	0.329	
	4.3.1	0.598	
	10.2.1	0.178	
Evaluator D vs Evaluator G	2.2.3	0.775	0.469
	6.2.3	0.492	
	9.8.5	0.619	
	11.4.1.1	0.317	
	4.3.1	0.528	
	10.2.1	0.084	
Evaluator D vs Output1	2.2.3	0.325	0.186
	6.2.3	0.245	
	9.8.5	0.192	
	11.4.1.1	0.067	
	4.3.1	0.115	
	10.2.1	0.175	

Table: Kappa score results (continued)

Comparison	Section #	K	Average K
Evaluator D vs Output2	2.2.3	0.663	0.294
	6.2.3	0.245	
	9.8.5	0.550	
	11.4.1.1	0.157	
	4.3.1	0.099	
	10.2.1	0.049	
Evaluator E vs Evaluator F	2.2.3	0.697	0.369
	6.2.3	0.495	
	9.8.5	0.358	
	11.4.1.1	0.292	
	4.3.1	0.239	
	10.2.1	0.131	
Evaluator E vs Evaluator G	2.2.3	0.697	0.382
	6.2.3	0.495	
	9.8.5	0.358	
	11.4.1.1	0.295	
	4.3.1	0.374	
	10.2.1	0.072	
Evaluator E vs Output1	2.2.3	0.438	0.194
	6.2.3	0.196	
	9.8.5	0.293	
	11.4.1.1	0.103	
	4.3.1	0.038	
	10.2.1	0.093	
Evaluator E vs Output2	2.2.3	0.596	0.304
	6.2.3	0.497	
	9.8.5	0.450	
	11.4.1.1	0.219	
	4.3.1	0.025	
	10.2.1	0.040	
Evaluator F vs Evaluator G	2.2.3	0.633	0.566
	6.2.3	1.000	
	9.8.5	0.542	
	11.4.1.1	0.239	
	4.3.1	0.554	
	10.2.1	0.426	

Table: Kappa score results (continued)

Comparison	Section #	K	Average K
Evaluator F vs Output1	2.2.3	0.267	0.172
	6.2.3	0.293	
	9.8.5	0.149	
	11.4.1.1	0.129	
	4.3.1	0.103	
	10.2.1	0.090	
Evaluator F vs Output2	2.2.3	0.542	0.239
	6.2.3	0.293	
	9.8.5	0.304	
	11.4.1.1	0.095	
	4.3.1	0.060	
	10.2.1	0.140	
Evaluator G vs Output1	2.2.3	0.394	0.213
	6.2.3	0.293	
	9.8.5	0.226	
	11.4.1.1	0.153	
	4.3.1	0.148	
	10.2.1	0.063	
Evaluator G vs Output2	2.2.3	0.697	0.376
	6.2.3	0.293	
	9.8.5	0.542	
	11.4.1.1	0.513	
	4.3.1	0.080	
	10.2.1	0.129	

Appendix F. Gold Standard

Ser.	Section #	Active Concept	Relation	Passive Concept
1	2.2.3	The Owner	shall furnish	a legal description of the site
2	2.2.3	The Contractor	shall be entitled to rely on	the accuracy of information furnished by the Owner
3	2.2.3	The Contractor	shall exercise	proper precautions relating to the safe performance of the Work
4	2.2.3	Owners	shall furnish	surveys
5	2.2.3	Surveys	describing	physical characteristics
6	2.2.3	Surveys	describing	legal limitations
7	2.2.3	Surveys	describing	utility locations
8	2.2.3	information	furnished by	Owner
9	2.2.3	precautions	relating	safe performance
10	6.2.3	The Owner	shall be reimbursed by	the Contractor
11	6.2.3	owner	shall be responsible to	the contractor
12	6.2.3	costs	incurred	by the owner
13	6.2.3	costs	are payable	separate contractor
14	6.2.3	costs	incurred	the contractor
15	9.8.5	The Certificate of Substantial Completion	shall be submitted to	the Owner for their written acceptance of responsibilities assigned to them in such Certificate
16	9.8.5	The Certificate of Substantial Completion	shall be submitted to	the Contractor for their written acceptance of responsibilities assigned to them in such Certificate
17	9.8.5	the Owner	shall make	payment of retainage
18	9.8.5	retainage	applying to	such Work
19	9.8.5	retainage	applying to	designated portion thereof
20	9.8.5	responsibilities	assigned	to them
21	9.8.5	payment	shall be adjusted	for work
22	9.8.5	work that	is	incomplete
23	9.8.5	work that	is not	in accordance with requirements
24	11.4.1.1	Property insurance	shall be on	an "all-risk"
25	11.4.1.1	Property insurance	shall be on	an equivalent policy form
26	11.4.1.1	Property insurance	shall include, without limitation	insurance against the perils of fire
27	11.4.1.1	Property insurance	shall include, without limitation	physical loss

Table: Gold Standard (continued)

Ser.	Section #	Active Concept	Relation	Passive Concept
28	11.4.1.1	Property insurance	shall include, without limitation	damage
29	11.4.1.1	damage	including, without duplication of coverage	theft
30	11.4.1.1	damage	including, without duplication of coverage	vandalism
31	11.4.1.1	damage	including, without duplication of coverage	malicious mischief
32	11.4.1.1	damage	including, without duplication of coverage	collapse
33	11.4.1.1	damage	including, without duplication of coverage	earthquake
34	11.4.1.1	damage	including, without duplication of coverage	flood
35	11.4.1.1	damage	including, without duplication of coverage	windstorm
36	11.4.1.1	damage	including, without duplication of coverage	falsework
37	11.4.1.1	damage	including, without duplication of coverage	testing and startup
38	11.4.1.1	damage	including, without duplication of coverage	temporary buildings
39	11.4.1.1	damage	including, without duplication of coverage	debris removal

Table: Gold Standard (continued)

Ser.	Section #	Active Concept	Relation	Passive Concept
40	11.4.1.1	Property insurance	shall cover	reasonable compensation for Architect's services
41	11.4.1.1	Property insurance	shall cover	reasonable compensation for Contractor's services
42	11.4.1.1	insurance	shall cover	expenses
43	11.4.1.1	debris removal	including	demolition
44	11.4.1.1	demolition	occasioned by	any legal requirements
45	11.4.1.1	expenses	required	insured loss
46	4.3.1	A Claim	is a	demand by one of the parties
47	4.3.1	A Claim	is a	assertion by one of the parties
48	4.3.1	the parties	seeking, as a matter of right	adjustment of Contract terms
49	4.3.1	the parties	seeking, as a matter of right	interpretation of Contract terms
50	4.3.1	The term "Claim"	includes	other disputes
51	4.3.1	The term "Claim"	includes	matters in question between the Owner and Contractor
52	4.3.1	matters in question between the Owner and Contractor	arising out of	the Contract
53	4.3.1	matters in question between the Owner and Contractor	relating to	the Contract
54	4.3.1	Claims	initiated by	written notice
55	4.3.1	The responsibility to substantiate Claims	shall rest with	the party making the Claim
56	4.3.1	the party	making	the claim
57	4.3.1	one of the parties	seeking	payment of money
58	4.3.1	one of the parties	seeking	extension of time
59	4.3.1	one of the parties	seeking	other relief with respect to the terms of the Contract
60	4.3.1	disputes in question between the Owner & Contractor	arising out of	the Contract
61	10.2.1	The contractor	shall take	precautions
62	10.2.1	Contractor	shall provide	reasonable protection
63	10.2.1	protection	to prevent	damage
64	10.2.1	protection	to prevent	injury
65	10.2.1	protection	to prevent	loss
66	10.2.1	utilities	not designated	for removal

Table: Gold Standard (continued)

Ser.	Section #	Active Concept	Relation	Passive Concept
67	10.2.1	utilities	not designated	for relocation
68	10.2.1	utilities	not designated	for replacement
69	10.2.1	structures	not designated	for removal in the course of construction
70	10.2.1	structures	not designated	for relocation in the course of construction
71	10.2.1	structures	not designated	for replacement in the course of construction

Appendix G. Precision and Recall Values

Evaluator	Section #	Precision		Recall	
A	2.2.3	1.000	0.834	0.444	0.615
	6.2.3	0.889		0.800	
	9.8.5	1.000		0.667	
	11.4.1.1	0.633		0.864	
	4.3.1	1.000		0.733	
	10.2.1	0.484		0.182	
B	2.2.3	0.900	0.561	1.000	0.858
	6.2.3	0.417		1.000	
	9.8.5	0.692		1.000	
	11.4.1.1	0.720		0.818	
	4.3.1	0.409		0.600	
	10.2.1	0.229		0.727	
C	2.2.3	1.000	0.904	1.000	0.929
	6.2.3	1.000		1.000	
	9.8.5	1.000		0.778	
	11.4.1.1	0.905		0.864	
	4.3.1	0.933		0.933	
	10.2.1	0.588		1.000	
D	2.2.3	1.000	0.895	0.778	0.588
	6.2.3	0.571		0.800	
	9.8.5	0.800		0.444	
	11.4.1.1	1.000		0.318	
	4.3.1	1.000		0.733	
	10.2.1	1.000		0.455	
E	2.2.3	0.875	0.734	0.889	0.633
	6.2.3	1.000		1.000	
	9.8.5	1.000		0.556	
	11.4.1.1	0.615		0.364	
	4.3.1	0.444		0.533	
	10.2.1	0.467		0.455	
F	2.2.3	0.778	0.708	0.778	0.712
	6.2.3	1.000		1.000	
	9.8.5	0.750		0.556	
	11.4.1.1	0.556		0.273	
	4.3.1	0.769		0.667	
	10.2.1	0.393		1.000	
G	2.2.3	1.000	0.842	1.000	0.978
	6.2.3	1.000		1.000	
	9.8.5	1.000		1.000	
	11.4.1.1	1.000		1.000	
	4.3.1	0.867		0.867	
	10.2.1	0.186		1.000	

Table: Precision and Recall Values (continued)

<table>
<tr><th>Evaluator</th><th>Section #</th><th>Precision</th><th>Recall</th><th>Evaluator</th><th>Section #</th></tr>
<tr><td rowspan="6">Computer Output1</td><td>2.2.3</td><td>0.800</td><td rowspan="6">0.476</td><td>0.444</td><td rowspan="6">0.376</td></tr>
<tr><td>6.2.3</td><td>0.375</td><td>0.600</td></tr>
<tr><td>9.8.5</td><td>0.429</td><td>0.444</td></tr>
<tr><td>11.4.1.1</td><td>0.500</td><td>0.227</td></tr>
<tr><td>4.3.1</td><td>0.375</td><td>0.267</td></tr>
<tr><td>10.2.1</td><td>0.375</td><td>0.273</td></tr>
<tr><td rowspan="6">Computer Output2</td><td>2.2.3</td><td>0.875</td><td rowspan="6">0.698</td><td>0.778</td><td rowspan="6">0.665</td></tr>
<tr><td>6.2.3</td><td>0.750</td><td>0.600</td></tr>
<tr><td>9.8.5</td><td>0.667</td><td>0.667</td></tr>
<tr><td>11.4.1.1</td><td>0.595</td><td>0.955</td></tr>
<tr><td>4.3.1</td><td>0.545</td><td>0.267</td></tr>
<tr><td>10.2.1</td><td>0.757</td><td>0.727</td></tr>
</table>