2009

An Event-Ontology-Based Approach to Constructing Episodic Knowledge from Unstructured Text Documents

Ting-Peng Liang
National Sun Yat-Sen University, tpliang@mail.nsysu.edu.tw

Dongsong Zhang
University of Maryland - Baltimore County, zhangd@umbc.edu

Ming-Yu Lee
National Sun Yat-Sen University, m954020017@student.nsysu.edu.tw

AN EVENT-ONTOLOGY-BASED APPROACH TO CONSTRUCTING EPISODIC KNOWLEDGE FROM UNSTRUCTURED TEXT DOCUMENTS

Ting-Peng Liang
National Sun Yat-Sen University
and City University of Hong Kong
Kaohsiung 80424, Taiwan, R.O.C.
tpliang@mail.nsysu.edu.tw

Dongsong Zhang
University of Maryland, Baltimore County
1000 Hilltop Circle,
Baltimore, MD 21250, USA
zhangd@umbc.edu

Ming-Yu Lee
National Sun Yat-Sen University
No. 70, Lienhai Rd., Kaohsiung 80424, Taiwan, R.O.C.
m954020017@student.nsysu.edu.tw

Abstract

Document summarization is an important function for knowledge management when a digital library of text documents grows. It allows documents to be presented in a concise manner for easy reading and understanding. Traditionally, document summarization adopts sentence-based mechanisms that identify and extract key sentences from long documents and assemble them together. Although that approach is useful in providing an abstract of documents, it cannot extract the relationship or sequence of a set of related events (also called episodes). This paper proposes an event-oriented ontology approach to constructing episodic knowledge to facilitate the understanding of documents. We also empirically evaluated the proposed approach by using instruments developed based on Bloom's Taxonomy. The result reveals that the approach based on proposed event-oriented ontology outperformed the traditional text summarization approach in capturing conceptual and procedural knowledge, but the latter was still better in delivering factual knowledge.

Keywords: Document summarization, design science, event-oriented ontology, episodic knowledge, knowledge management

Introduction

With the rapid growth of the World Wide Web and electronic information services, digital information is increasing at an incredible rate, causing the unprecedented problem of information overload. No one has time to read everything, yet we often have to make critical decisions based on what we are able to assimilate. Thus effective management of electronic documents, especially management of complexity and specialization of knowledge expressed in those text documents, is essential to enterprise knowledge management. One challenge that managers face is how to construct deep knowledge from a collection of documents to support problem solving. For instance, given a large collection of documents about the financial tsunami, how can we use information technology to gain insights or to extract useful knowledge about this phenomenon from those documents so that we can handle it better in the future or prevent it from happening again? Without such capabilities, the value of a knowledge management system (KMS) would be limited to a static digital storage rather than a powerful decision aid. Developing such capabilities, however, is by no means an easy task.

The technology of automatic text summarization is one of the major tools indispensable for dealing with information overload. It is aimed to distill the most important information from a text document and produce an abridged version of the document for easy and quick grasp of its main idea. Most automated summarization systems today produce extraction based summaries, which uses a simple and language-independent summarization strategy by identifying the most important/topical/informative issues of the text and assembling them together. Although a summary is not necessarily coherent, people can still form an opinion about the original content. Different extraction strategies have been developed (e.g., Antiqueira et al., 2009; Hennig et al., 2008). These approaches are usually useful in extracting sentences from documents that may not have close relationships or focus on presenting factual information. However, they are inadequate for presenting relationship between certain events in a complex domain. For example, it is natural to examine the previous government and banks' reactions in the 1929 recession when the financial crisis appeared in 2008 to help determine what should be done. In order to discover those reactions of different organizations from historical documents, a decision maker will need not only summaries of those reports but also a clear illustration of the unique events, their sequential relationships, as well as roles involved in those events. A set of events and their sequential relationships in a certain time period is called an *episode* (Mannila et al., 1997). To make a KMS useful, the ability to discover episodic knowledge is definitely a challenging but critical function. The current text summarization techniques, however, fail to achieve that ability.

Ontology is one of the fundamental cornerstones of knowledge management, as well as the building block of the Semantic Web. An ontology can be defined as a formal, explicit, and shared conceptualization of the domain of discourse that defines concepts and relationships within the domain (Felden and Kilimann, 2006) and demonstrates the knowledge structure of the domain (Gruber, 1993). Traditionally, ontologies are often described using knowledge representation techniques, such as frame and predicate logics. Concepts in an ontology form a class hierarchy, and subclasses inherit properties of superclasses. Structurally, an ontology is a graph whose nodes and arcs represent conceptualizations, independently from how to assign formal semantics to these conceptualizations. Ontology provides not only a semantic ground for machine-understandable description of digital content, but also a common layer that plays a major role in supporting information exchange and sharing by extending syntactic interoperability to semantic interoperability (Karoui et al., 2006).

This study is aimed to make multi-fold contributions to the development, presentation, and evaluation of an event-oriented ontology based approach to constructing episodic knowledge from a collection of news documents. We also empirically evaluate the value of the proposed approach in facilitating users' learning and understanding, in comparison to the traditional text summarization approach. The remainder of the paper is organized as follows. Section 2 introduces related work on ontology development and document summarization. We will introduce the proposed event-oriented ontology approach in Section 3. Section 4 presents the method of empirical evaluation and results, followed by discussion and conclusion in Section 5.

Related Work

Document Summarization and Episodic Knowledge Discovery

There has been extensive research on text summarization. In general, there are two approaches to text summarization: knowledge poor and knowledge rich approaches. The former tries to evaluate the importance of a sentence in a document by using some weighted features, such as the frequency of words, title words, cue words/phrases, the location of sentences, and the syntactic structure of sentences. The sentences with the highest scores are regarded as the most significant and then extracted. This approach does not perform any semantic-level analysis and usually does not require deep knowledge. In contrast, the knowledge rich approach tries to analyze a text document using knowledge, such as the grammar or lexical databases of the target language. This approach relies on a priori built-in knowledge. It is usually domain specific and more complex due to the difficulties in building an effective machine usable knowledge base.

In terms of the source of text summarization, there are approaches to single-document summarization and to multi-document summarization. Single-document summarization is essential to enabling and improving quick access to large quantities of information. Recently, CNN.com added "Story Highlights" to many news articles on its site by giving a brief overview of the article with three or four related sentences in the form of bullet points, aiming to allow readers to quickly gather information about those stories. Multi-document summarization identifies and synthesizes important phrases or sentences across a number of documents that address the same topic (Barzilay et al., 1999, Nastase 2008). Examples of multi-document summarization systems include SUMMONS (McKeown and Radev, 1995) and NeATS (Lin and Hovy, 2002). The former extracts important information from different documents by instantiating slots of a set of pre-defined templates to summarize a series of news articles reporting the same event, while the latter is an extraction-based multi-document summarization system. It leverages techniques proven effective in single document summarization such as term frequency, sentence position, and stigma words to select and filter content.

As introduced earlier, most existing methods for both single- and multi-document summarization rely on content extraction (e.g., sentence and paragraph extraction), which is a knowledge poor approach. With extraction based summarization, key sentences in documents need to be identified and ranked based on their occurrence frequency as well as the position of their appearance in the document. The most important sentences are then used to construct a summary. However, text summarization does not reveal the semantic relationships among the concepts, entities, roles, and actions reported in a document. An early work by Mannila et al. (1997) identified the need to discover episodic information in documents. In addition to identifying similarities among documents, the work by Mani and Bloedorn (1999) selected important differences in individual documents to summarize a set of news reports about an event or a sequence of events. Although those early works pointed out the importance of episodic information, they did not take advantage of the ontological information in their analysis. With the recent advancement of ontology development, we are able to improve the methodology for constructing episodic knowledge.

Ontology Development

There are two main approaches to facilitating ontology construction from a language processing perspective. The first approach helps manual ontology engineering by providing natural language processing tools to support shared decisions and ontology import. It involves interviewing experts, transcribing into text, and manually analyzing the text to identify object-attribute pairs that can be incorporated in the ontology. The second approach relies on machine learning and automated language processing techniques to extract concepts and ontological relations from structured and unstructured data such as databases and text (Navigli et al., 2003).

A number of methods for gaining and modelling knowledge in ontologies have been proposed. OnToKnowledge (Davies et al., 2002) is based on software engineering lifecycle models, starting from requirement analysis to the maintenance of the developed ontology; the Skeletal methodology (Uschold and King, 1995) comprises a set of guidelines for developing ontologies; methontology (Blázquez et al., 1998) supports development-oriented activities, and describes project management activities; OntoClean (Guarino and Welty, 2004) assesses the ontological adequacy of taxonomic links in ontologies; Rapid Ontology Development (ROD) method (Zhou, 2007) consists of three phases: design, learning, and validation. The design phase involves the identification and detailed analysis of domains, requirements, and relevant resources with the help of users and/or domain experts. The output includes specifications of domains, intended applications of ontologies, and authoritative domain sources. In the learning phase, appropriate learning techniques are selected, implemented, and then applied to discover ontologies from domain sources; the learning results are then evaluated and the developed ontology will be refined during the validation phase, where the created ontologies are checked for errors, redundancy, conflict, and comprehensiveness.

In general, to construct an ontology, specialists must thoroughly analyze a domain by (Navigli et al., 2003):

- term identification: creating a vocabulary that describes the entities that populate the ontology by information extraction (i.e., concept identification);
- developing formal descriptions of the terms in that vocabulary; and
- characterizing the conceptual relations among those terms.

1) Identification of important terms in a domain

The first step of creating a vocabulary for an ontology is to extract important terms from text documents related to a particular domain. Terminology can be considered the surface appearance of important domain concepts. Candidate terms are usually captured with shallow processing techniques that range from stochastic methods to more sophisticated syntactic approaches (Navigli et al., 2003). The richer the syntactic information is, the higher the quality of the result will be.

Generally, a collection of documents (called corpus) is used as input to a text mining algorithm. The corpus is then parsed into tokens (i.e., contiguous string of characters delimited by space, punctuation, or other character separators) or terms (tokens in a particular language). The unstructured text in the corpus becomes a structured data object via the creation of a term-by-document frequency matrix, to which numerical measures can be used to weigh terms (Inniss et al., 2006). Since the overall goal is to develop an ontology, a straightforward approach is to find those salient concepts that occur most in the majority of documents in the collection. High occurrence frequency in a corpus is a property observable for terminological as well as non-terminological expressions. Frequency weights of those concepts can be adjusted to account for the distribution of terms across documents (e.g., using Entropy or Inverse Document Frequency (IDF)).

It is well recognized that natural language processing (NLP) and text mining techniques are effective for information extraction from text documents. Some NLP tasks related to ontology development are information extraction and automatic summarization (Eom and Zhang, 2004, Velardi et al., 2001).

2) Formal representation of terms

Ontology representation is fundamental in ontology development. In addition to making ontologies understandable by computers and users, an ontology representation language should also provide representation adequacy and inference efficiency. The standardization of ontology representation languages (e.g., Web Ontology Language (OWL)) has taken big strides in the past few years. Some languages have adopted a frame-based knowledge representation paradigm, while others incorporate description logics to enhance the expressiveness of reasoning systems (e.g., Stevens et al., 2002).

3) Semantic interpretation for identifying term-relationship

This step includes semantic interpretation of identified individual terms/concepts and linking them within a domain based on their semantic relationships to form a graph. In the simpliest way, relationships between a pair of terms in a domain can be identified based on the frequency of term co-occurrance in the same documents. However, such relationships based on co-occurrance frequency are shallow and do not reflect any semantic meaning. Generic ontologies such as WordNet and HowNet (www.keenage.com) can be employed to find the semantic meaning of the hidden relations and patterns, although they might be too general to describe domain specific knowledge. Therefore, combining the strength of both generic ontologies and machine learning methods while attempting to find the hidden relations and patterns from documents seems more effective (Yang et al., 2004, Navigli et al., 2003).

It is worth pointing out that prior studies on ontology learning generate ontologies that can't demonstrate the dynamic and evolving nature of events in a specific domain. For example, although an ontology may include individual persons, concepts, and activities involved in related events, it fails to answer questions about the roles of individuals in various activities in a certain event, as well as the sequence of activities occurred in the event.

An Event-oriented Ontology Approach

In his classic paper, Alan Newell characterized knowledge as a behavioral phenomenon (Newell, 1982). He viewed knowledge in terms of goals of an agent, the actions of which the agent might be capable, and the means by which the agent selects actions in order to achieve its goals. This view of knowledge goes well beyond the notions of conceptualization specification and of an enumeration of concepts and relationships (Brewster and O'Hara, 2004). From Newell's perspective, knowledge directly connects goals with actions. In that sense, knowledge has a strong procedural element. However, most existing ontologies do not contain or reflect such a 'procedural element' of knowledge, which is very important for problem solving.

To fill the void in the literature, we propose a novel approach that constructs an event-oriented ontology from a collection of text documents. The proposed method retrieves event-related concepts from documents, identifies relationships among episodes, events, and sub-events, establishes the linkage between roles and activities in a temporal manner, and provides an interactive visualization tool to facilitate users in editing and presenting the ontology. The research follows the design science research framework proposed by Hevner et al. (2004), including an experimental evaluation of the effectiveness of the proposed methods.

Design and Development of Event-oriented Ontology

There are a variety of challenges in extracting and processing events reported in news. First, it is common that a piece of news may include multiple topics or events, or one event can be covered by several news reports, so considering one piece of news as one topic or event is inappropriate. An event-oriented ontology, which we refer to as an ontology that focuses on the events related to a certain topic, relationships (e.g., procedural and temporal) among those events, as well as roles involved in those events, has to be able to reflect such many-to-many relationships. Second, any event has to be described from multiple dimensions in the ontology, such as roles involved, time occurred, and what happened. Stimulated by the idea of OLAP (Online Analytical Processing), we propose an approach to event-oriented ontology construction by using a hierarchical "sub-event→event→topic" ontology structure, coupled with a semantic relation repository, similarity comparison, and merge of sub-events, to record many-to-many relationships. Such an ontology allows different views (e.g., role, incident, and time) of the same event from various angles and from a temporal perspective. Users can easily see the sequence of related events and corresponding time periods through a visualization tool, which can lead to a quick grasp of an occurred event.
The proposed method for creating an event-oriented ontology, as shown in Figure 1, consists of three sub-systems: pre-processing documents, ontology construction, and displaying ontology.

Pre-processing sub-system

The major objective of pre-processing sub-system is to extract important concepts about the same event from text documents automatically (e.g., Chinese news reports in this case), then apply Apriori association rule mining algorithm (Agrawal and Srikant, 1994) to discover relationships between concepts. This phase includes POS (Part-of-speech) tagging, word filtering, and term analysis.

In our study, we used a Chinese natural language processing tool called CKIP to process news documents by marking syntactic annotations and segmenting words and sentences. This process is facilitated by specialized named-entity and location dictionaries. Then the word filtering step removes non-important terms, including stop words, and mainly keeps nouns, verbs, and phrases. Finally, the term analysis will help identify important terms and determine their potential relationships. In this step, TF-IDF (Term Frequency/Inverse Document Frequency) weight is calculated for each individual term. Then, all the terms are ranked based on their TF-IDF weights. Those with smaller weight values are considered less important to the event and therefore will be discarded. The Apriori algorithm will be applied to the remaining terms to discover the potential relationships between them. Only those relationships that have support and confidence values higher than pre-defined thresholds will be considered important. At the end, the major concepts in documents, including person, tasks (i.e., actions), time, location, objects, and their relationships identified by the Apriori algorithm, are stored in a lexical database.

Ontology construction sub-system

This sub-system provides a friendly and interactive event ontology editor. Ontology engineers can edit concepts, attributes, and relationships, and construct people, task (i.e., action), location, and object ontologies.

The content of a piece of news may consist of several roles and actions performed by those roles. We call such subject-verb-object combinations as sub-events. Those sub-events are assembled together in order to describe an issue. However, they could belong to different event categories. Therefore, the proposed method uses a bottom-up approach to select, re-organize, and classify sub-events distributed in different news. It starts with identification of sub-events, then merges similar sub-events into clusters and defines those clusters as events, and finally forms related events into topics and eventually into a case to build a comprehensive event ontology, as shown in Figure 2. The procedures for building the event-oriented ontology include the following six steps, which can be described in a simplified example below.

[News #1] *The firefighters arrived at the scene within 10 minutes. One person died and eleven others were injured. They were intoxicated by carbon monoxide. The firefighters measured the gas density. The mayor, Ma Ying-Jiou, expressed his condolences and asked for better safety regulations.*

[News #2] *Mr. Dexiong Liang died in the Alexander accident. The city government apologized and called a meeting today. The government ordered to inspect all health clubs. Yachun Tang, CEO of Alexander, apologized to the public but she also blamed government policy problems for the accident.*

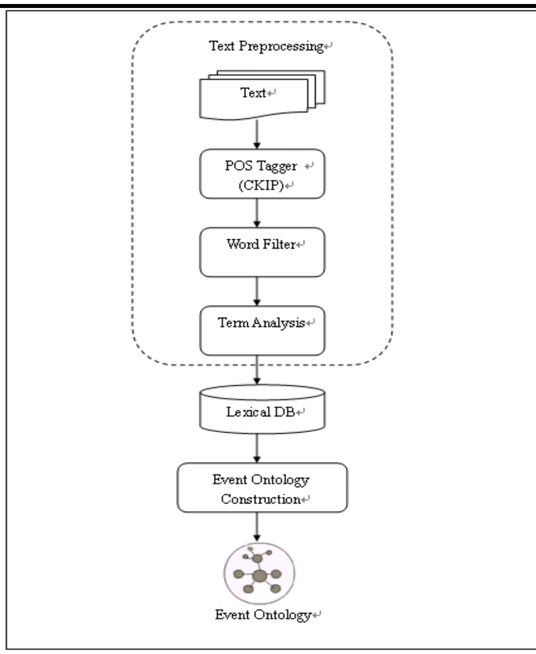

Figure 1. The Methodology for Event Ontology Construction

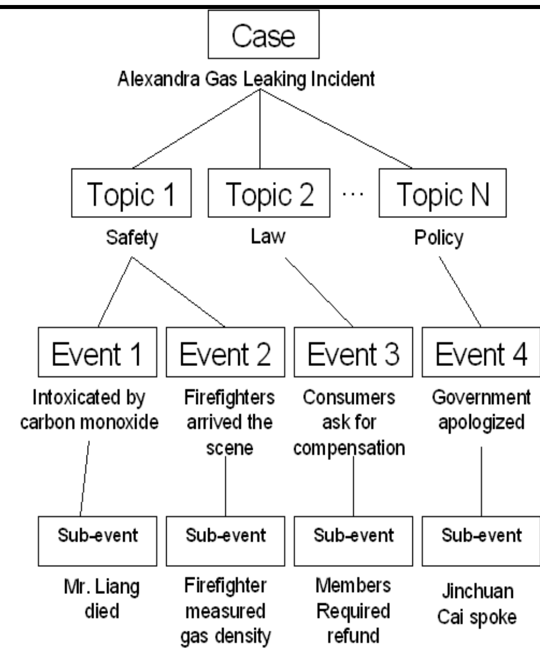

Figure 2. Converting News Documents to An Event-oriented Ontology

Step 1: Creating sub-events

A sub-event is considered as a simple sentence consisting of a subject-verb-object structure. The proposed method selects important relationships between nouns and verbs identified by the Apriori algorithm during preprocessing, and uses those important verbs as the core of sub-events to find associated subjects and objects for constructing sub-events. For instance, "firefighters measured gas density" is a sub-event.

Step 2: Creating a noun-noun relationship lexicon

This step selects those terms that have strong noun-noun relationship identified by the Apriori algorithm and stores them in a relationship lexicon for step 3. For example, Ma Ying-Jiou and City Mayor are defined as two related terms (is-a relationship) in the sample document. This allows other sub-events related to the mayor and Ma Ying-Jiou to be linked easily.

Step 3: Similarity comparison

The purpose of similarity comparison is to determine the level of similarity among sub-events. If the similarity level between two sub-events is high, it implies that either sub-events may describe the same event and therefore should be merged into one sub-event, or their contents are similar, related, and should be linked to the same event. In this research, we extended similarity measure refined by Xu (2002) by incorporating relationship lexicon generated in step 2. For example, the sub-event "One person died" in the first document and the sub-event "Dexiong Liang died" in the second are considered to be similar events.

Step 4: Merging sub-events

After getting the similarity matrix of sub-events in the previous step, this step will select and merge sub-events that have similarities higher than a pre-defined threshold. Because each sub-event extracted from a news document has a temporal point, the system determines whether or not two sub-events should be merged into one sub-event based on their occurrence time and content. For instance, the similar sub-events concerning the death of Dexiong Liang can be merged to become a single sub-event "Dexiong Liang died."

Step 5: Editing sub-events

It consists of two parts: editing attributes of sub-events and editing relationships of key concepts. Each sub-event can be considered as a concept. Concept attributes include person (role), time, location, and object, etc. This approach automatically extracts the results of pre-processing sub-system and presents the structure of each event. Users can edit or modify the key terms automatically categorized as person, location, time, and object by the system. In addition, ontology engineers can define three relationships here based on the previous results of the Apriori algorithm: association, aggregation (IS-PART-OF relationship), and generalization (i.e., IS-A relationship). This allows the sub-events related to the City Mayor and those related to the City Government to be grouped.

Step 6: Editing person, location, and object ontologies

The event ontology is constructed using the bottom-up approach: from sub-event to event, and eventually to topic. Each sub-event in the generated ontology is represented by the combination of actions and other attributes, as shown in Figure 3. The system will also generate person, location, and object ontologies separately using the same bottom-up approach. The system provides information to help users construct the ontology, such as the time when events occurred and information about other related sub-events while processing one sub-event.

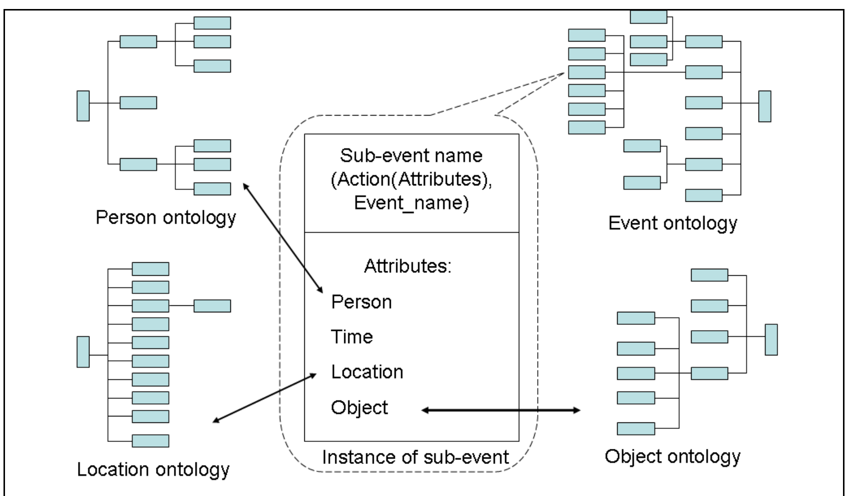

Figure 3. An Sub-event Instance

Displaying Ontology: Presentation of Episodes

Traditional ontology is difficult to read, especially when it is illustrating the relationships among multiple events (episodes). A formal, graphical representation not only is more understandable, but also provides a much more consistent vehicle for conveying ontological concepts and for sharing them with other domain experts not versed in the representation language (Ceccaroni and Kendall, 2003).

Our prototype system has a sub-system to provide an interactive ontology visualization tool that enables users to easily navigate and understand people, action, place, and their relationships in the ontology. In addition, users can also understand the context and scenario of an event through a flow chart. Figure 4 shows a sample flow chart of a summarized episode. An episode can be viewed in two ways: exploring scenarios of events and sub-events from a topic (Figure 4a), or viewing its sub-events based on the roles participated (Figure 4b). In Figure 4a, the oval TO_i represents a selected topic; rectangles E_1, E_2, … E_6 represent events; the left-to-right arrow indicates the time sequence. The events in the same column occur at the same time, while different columns refer to different time points. In Figure 4b, the oval represents an event, and small circles underneath are roles involved in the event E_k; rectangles represent sub-events; matching colors of roles and sub-events indicate the sub-events in which a role is involved. Figure 5 in Appendix 1 shows two events that the Taipei City Government involved in the Alexander case.

The use of these two views can be explained below. If we want to know how many legal issues are involved in the financial tsunami (i.e., an event), then Figure 4a can show what actions he took in different occasions over time if we choose the legal topic. If we want to view how president Obama and CEO of AIG did in the AIG bailout case, then Figure 4(b) would be appropriate, where U1 and U2 represent Obama and the CEO of AIG, respectively, and S_i represents a sub-event associated with either U1 or U2. A user can click any item in the chart to view the details in original news that are related to the selected item.

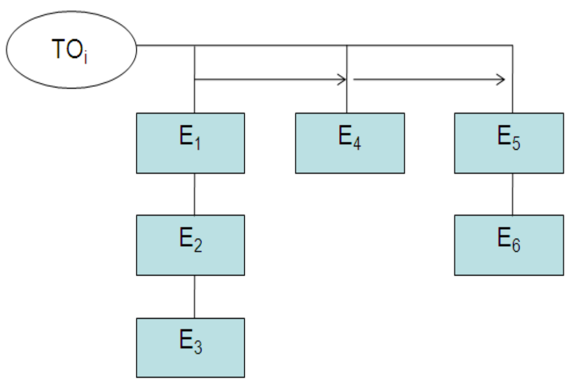

Figure 4a. An Event Flow Chart with Topic as the Main Axis

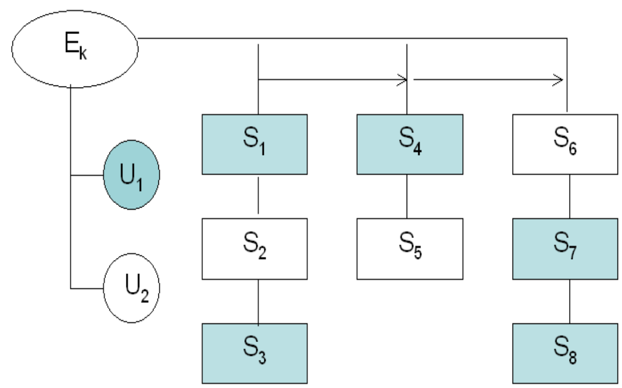

Figure 4b. A Sub-event Flow Chart with Event as the Main Axis

Based on the above system design, we implemented the proposed approach in a prototype system (i.e., artifact), which was used in an empirical evaluation study, which will be introduced in the following section.

Empirical Evaluation

Evaluation Appraoch

There are normally two different views of the assessment of information systems: one is to focus on evaluating system functions; the other focuses on the system facilitation in understanding information. Specifically, the quality of an ontology may be assessed in various dimensions, such as structural (e.g., the topological and logical properties of an ontology), functional (e.g., functions and design of ontology), and usability-related perspectives (e.g., ease-of-use of ontology).

There are inherent problems in evaluating an ontology as it is not clear what exactly one is trying to evaluate. Many researchers used the notions of *precision* and *recall* that are commonly used in the evaluation of information retrieval or classic natural language processing systems to evaluate an ontology (e.g., Vargas-Vera and Celjuska, 2004). Precision tries to measure the amount of knowledge correctly identified (in the ontology) with respect to the whole knowledge available in the ontology. Recall reflects the amount of knowledge correctly identified with respect to all the knowledge that should be identified. Brewster et al. (2004), however, suggest that ontology evaluation cannot be compared to those evaluation tasks. They argue that an ontology is a representation or model of knowledge. As Gruber defined, ontology is a “formal, explicit specification of a shared conceptualization” (Gruber, 1993). The word ‘shared’ implies that a developed ontology may be extremely subjective, representing the time, place, and cultural environment in which it is created. Precision and recall depend on a clear set of items concerned. There is no clear set of “knowledge to be acquired” because the same set of facts can give rise to very different interpretations and therefore different kinds of knowledge. Therefore, Brewster et al. suggest that precision and recall measures are not appropriate for evaluating ontology. Other explored metrics include cost-based evaluation metric (e.g., error rate), ontology fit (i.e., measuring the “fit” between an ontology and domain of knowledge), and usefulness and/or relevance for practice (Hartmann et al., 2004). Because the focus of our evaluation is on whether the proposed event-oriented ontology can facilitate in users’ understanding and learning about news events, so the cost-based and ontology fit metrics are not appropriate either.

The problem of qualitative approaches to evaluating an ontology by presenting users with the ontology and asking them to rate it lies in how to determine who the right users are, and what criteria to provide them for their evaluation. One standard approach is to compare an ontology generated by the proposed approach with an existing ‘gold standard’, or with an ontology generated by experts manually (e.g., Inniss et al., 2006). The problem is that if the results differ from the gold standard, it is difficult to determine whether that is because the used corpus or methodology is inappropriate, or whether there is a real difference in the knowledge present in the corpus and the gold standard (Brewster et al., 2004).

Typically, an ontology will be used in some kind of application or task. The outputs of the application, or its performance on the given task, might be better or worse depending partially on the ontology used. Therefore, another potential effective approach to ontology evaluation would be to evaluate how effective a particular ontology is in the context of an application (Brewster et al., 2004). This is reasonable in the sense that a relatively straightforward and non-problematic evaluation approach may already exist for the output of the application (Brank et al., 2005).

Based on the above pros and cons, we decided to conduct the evaluation of the proposed event-oriented ontology in the context of a real world task, namely information search. We used a document summarization system as the benchmark. Goodman (1973) believes that reading is not simply obtaining meanings from words, but a process of meaning construction by organizing relationships of existing concepts. Understanding is the ultimate goal of reading, so are summaries. When searching for news or information about a certain event from a digital archive, presenting summaries of news is often very beneficial to users. Therefore, it is more interesting to examine if users supported by an event-oriented ontology can improve document reading and understanding in comparison with those supported by a text summarization system than to evaluate technical functions of the prototype system directly.

Theoretical Foundation

We used Bloom’s Taxonomy of Educational Objectives (Bloom, 1969) to develop instruments for measuring the effectiveness of users’ learning. Bloom’s taxonomy is well-accepted in educational research to measure different levels of learning goals. The theory proposes that the cognitive domain involves knowledge and the development of intellectual skills, which include recall or recognition of specific facts, procedural patterns, and concepts that serve in the development of intellectual abilities and skills. The model includes six major categories, starting from the simplest behavior to the most complex:

1. **Knowledge** of terminology; specific facts; ways and means of dealing with specifics; universals and abstractions in a field (principles and generalizations, theories and structures). Knowledge is defined as the remembering (recalling) of appropriate, previously learned information.
2. **Comprehension**: Grasping (understanding) the meaning of informational materials.

3. **Application**: The use of previously learned information in new and concrete situations to solve problems that have single or best answers.
4. **Analysis**: Breaking down of informational materials into component parts, examining (and trying to understand the organizational structure of) such information to develop divergent conclusions by identifying motives or causes, making inferences, and/or finding evidence to support generalizations.
5. **Synthesis**: Building a structure or pattern from diverse elements. Putting parts together to form a whole, with emphasis on creating a new meaning or structure.
6. **Evaluation**: Judging the value of material based on personal values/opinions, resulting in an end product, with a given purpose, without real right or wrong answers.

Because Bloom's taxonomy of cognitive domain fits the assessment of document understanding very well, we decided to use it (mainly knowledge and comprehension dimensions) as the theoretical foundation for evaluating users' cognitive understanding of documents using the proposed ontology.

Experimental Systems

In the empirical evaluation, for simplicity, we developed a two-level news summarization system that provides summaries for a collection of news, as shown in Figure 6 in Appendix I. At a higher level, the system extracted the title of every piece of news and presented all the titles in a chronological order in the left panel of system interface as the summary of entire news collection. If a user is interested in a specific piece of news after reading this general summary, he/she could simply click the title of that news to view its summary (single news summary) in the right panel of the system interface. The user could further press the 'Full article' button located at the right upper corner of the system interface to view the full content of the news.

The lower-level single-document summarization sub-system consisted of three parts: news analysis, term weighting, and automatic summarization. The news analysis segments sentences in a Chinese news document; then in the term weighting part, Chinese terms in sentences would be separated by a natural language processing tool called CLIP (http://ckipsvr.iis.sinica.edu.tw) to identify key terms and part-of-speech tags. We mainly kept noun phrases (e.g., noun, noun-verb, and noun-noun) because the key terms in a document are mostly noun phrases. We also kept the verbs because they represent the relationship between noun phrases. Considering long terms normally have lower occurrence frequency in a document in comparison to short terms but carry more unique meanings, we took the number of words in a term (i.e., the term length) into account when calculating term weight. Longer terms (e.g., management information systems) have higher priorities to be selected than shorter terms (e.g., management, information or information systems).

Finally, in the automatic summarization part, the first step is to find the most important sentence(s). We used Jaccard method to assess the similarity between any two sentences based on the occurrence of the same terms. The total similarity score of one sentence is the sum of all similarity scores between this sentence and other sentences in the document. After calculating total similarity scores of all sentences in a document, we chose the top five sentences with the highest similarity scores and assembled them in the sequence of their occurrence in the current news as the summary. Appendix I shows the screenshots of both ontology system and two-level summarization system interfaces used in the evaluation.

Experiment

Participants: We recruited participants from Management Information Systems major at a large university in Taiwan for this study. All volunteered participants were interviewed right before the study to examine the level of their knowledge about the events reported in news that would be used for the evaluation to ensure that participants knew little about them in advance. This pre-screening was done by asking participants a few randomly chosen questions related to those events. Eight participants were excluded from the study because of their high a-priori familiarity with those events. Finally, sixty undergraduate (28) and master students (32) qualified for and participated in the experiment. Every participant had years of experience with computer and the average weekly computer usage time was at least 20 hours. Among all participants, 44 were male; 90% of participants were 18~25 years of age and the rest were between 26 and 30 years of age.

Cases used in the experiment: Two news cases were used in this experiment. One was the Alexander gas leaking case (We call it 'Alexander case'). Alexander Health Club is a well-known health-training center in Taiwan. On January 31st, 2006, one of its branches in Taipei had a serious carbon monoxide leaking accident. There were twelve club members and employees intoxicated. After delivered to the emergency room, one club member, Mr. Dexiong Liang, died because of inhaling a large volume of carbon monoxide. Other eleven people were recovered after the medical treatment. That incident raised considerable concerns from government and public about the safety of health clubs. Local governments explored a variety of actions, laws, and policies in order to improve safety and prevent similar tragedy from happening again. The Alexander Health Club apologized to the public and closed the business for a number of internal revision and safety enhancement. It did not reopen the business until the government agencies ensured that it passed the safety requirement. This case consisted of twenty-seven pieces of news on the occurrence of the accident, reactions of local government, security check, responses from Alexander, and discussion about public safety, etc. The lengths of the news reports varied from 300 to 800 words, with an average of 550 words.

Another case was the acquisition of the struggling Siemens mobile group by BenQ Corp., a Taiwan-based IT company, in 2005 (We call it 'Siemens case'). The goal of the acquisition was to combine BenQ's lifestyle experience and renowned design team with Siemens' engineering capabilities to create a new leader in the mobile communications market. Unfortunately, the German division of the new company filed for bankruptcy in a Munich court in 2006. Since then, BenQ didn't intend to continue manufacturing mobile phones in Germany. It included seventeen pieces of related news about the reasons of the acquisition and its failure, operations and reactions of BenQ, and actions of German Supreme Court, etc. The lengths of those reports varied from 200 to 1100 words, with an average of around 550.

Task: Each participant was randomly assigned to one of two groups, one only using the prototype event-oriented ontology system and another only using the benchmark summarization system. At the end, there were thirty participants in each group. After system training, participants were asked to 'browse' the news documents of the first case. Then, they were asked to answer questions regarding specific knowledge contained in the documents they just browsed (See question examples in Appendix II). The participants did not know the questions they had to answer before accessing the system. Once they finished, they would be given the second collection of news related to another case and repeated the same procedure. The sequence of Alexander and Siemens question sets was randomized. It is worth noting that in order to assess the potential benefits of the proposed ontology and summarization approach, we intentionally provided participants with a restricted time (ten minutes for each case) so that they would not be able to read every single document completely. Subjects in either group were given the same amount of time to complete the task.

Design of questions: Questions used in the experiment were developed based on the knowledge and comprehension dimensions of Blooms' taxonomy. Because both cases consisted of news, we selected factual, conceptual, and procedural knowledge in line with the knowledge dimension. Also because part of the definition of comprehension is to create new knowledge from old knowledge, which is not suitable for news events, so we chose knowledge recall measurement from the comprehension dimension. Therefore, the questions for each case consisted of three parts, with four questions in each part: questions about facts, about concepts, and about procedures reported in the case. Examples of questions used in formal experiment are provided in Appendix II.

After the questions were developed, we conducted a pilot study with ten participants (also university students but none of them participated in the formal experiment later). They were randomly assigned to use either even-oriented ontology system or the summarization system, and answered all three types of questions that would be used in the formal experiment. No participant ever reported any problem with either system in both pilot and formal experiment.

Data Analysis and Results

In this study, a participant's task performance was measured by his/her final score calculated based on how many questions were correctly answered. Each correct answer added one point and there was no partial credit for answers. Therefore, the maximum score of each participant for each knowledge part is four. Participants were allowed to use the system to navigate news content while answering questions. In order to motivate participants to accomplish the experimental task seriously, participants were informed that top performers who received the highest scores in the task would get monetary rewards.

Table 1 shows the means, standard deviations, and difference of participants' scores of two groups with both tasks. Results of a paired t-Test reveal that given limited navigation time, participants using the system supported by the proposed event-oriented ontology significantly outperformed those using the text summarization system in recalling concepts and procedures, but was outperformed by the benchmark system in recalling facts.

Table 1. Overall User Performances of Two Systems

Measures	**Event-oriented ontology System (O)**		**Text Summarization System (S)**		**Difference in means**	P
	Mean	**SD**	**Mean**	**SD**	**(O-S)**	
Recall of facts	2.150	0.458	3.867	0.225	-1.717	0.00**
Recall of concepts	3.683	0.334	2.417	0.373	1.266	0.00**
Recall of procedures	3.650	0.326	1.667	0.442	1.983	0.00**

Note: 1) The performance scores are ranged from 0 to 4, with 4 being the best score;
2) **: *P*<0.01

In order to minimize the potential difference in two news cases, we also analyzed data of two cases separately. The results shown in Table 2 are consistent with those in Table 1. As expected, the results clearly show that the user can better understand conceptual and procedural knowledge when they are provided with a system that presents synthesized information. The proposed ontology-based approach for organizing and presenting a collection of documents can improve user understanding of conceptual relationships and procedural knowledge embedded in the documents.

Table 2. User Performance of Two Systems in Two Cases

Measures	**Alexander**			**Siemens**		
	Event-oriented	**Summa-rization**	**P**	**Event-oriented**	**Summa-rization**	**P**
Recall of facts	2.133	3.967	.00**	2.167	3.766	0.00**
Recall of concepts	3.733	2.167	.00**	3.633	2.667	0.00**
Recall of procedures	3.633	1.767	.00**	3.667	1.567	0.00**

Note: 1) The performance scores are ranged from 0 to 4, with 4 being the best score;
2) **: *P*<0.01

Discussion and Conclusion

We have presented an event-oriented ontology method for constructing episodic knowledge. The empirical evaluation shows that the user can better capture conceptual and procedural knowledge while the text-based summarization performed better in capturing factual knowledge. The possible reason for this phenomenon is that the ontology is constructed based on concepts (e.g., person, objects, sub-event, event) and their relationship in related news events, which intuitively would help answer questions related to those concepts. In addition, the visualization tool provides a sequnetial view of various events, which demonstrates the temporal and procedural knowledge more explicitly and effectively than the summarization system. However, the current two-level summarization system allows a user, given a factual knwoledge question, quickly identify which news may contain the answer and then view that piece of news immediately, making it easier to locate the answer about factual knowledge than the event-oriented ontology system. It implies that different approaches (e.g., event-oriented ontology and text summarization systems) should be used to deal with different knowledge inquiries. In particular, event-oriented ontology can better support problem solving that requires more complex and procedural knowledge.

This research has multi-fold contributions. First, we propose to use person, action, location, and object aspects to organize and describe an event, and use time to provide sequential relationships betweeen activities involved in an event. The higherachical structure of event-oriented ontology can help capture many-to-many relationships between documents and events, and help anwer specific questions related to an event. Such an ontology enables users to quickly grasp the gist of a series of related news archived in a digital library.

Second, we propose an effective method for building event ontologies. We use a pre-processing subsystem to extract key terms and their relationships automatically; then adopt a bottom-up approach to create, classify, and merge sub-events into events, which are further grouped into different topics. In the meantime, the method also creates separate ontologies for person, location, and objects.

Third, the generated event ontology can be displayed in a visual format through flow charts, which can better facilitate users to understand the entire scenario of events quickly.

Fourth, there has been extensive research on ontology development in the literature, but relatively little effort on ontology evaluation from a user perspective. In this study, we developed a theory-based evaluation instrument for evaluating the effectiveness of the proposed approach in knowledge management systems. The instrument is based on the well-known Broom's taxonomy that allows the researcher to assess different learning effects.

There are a few limitations of this research that merits further investigation. First, we used news documents only in this study. It would be interesting to examine if similar findings can be obtained with other types of documents. Second, the text summarization system used in this study is appropriate for long articles because it extracts and assembles most important sentences appeared in documents. Therefore, the longer the article, the more precise the summary. Becausee news articles are relatively short, the generated summaries may or may not be in the best qualilty. We plan to fine-tune the system by incorporating new heuristics so that it can handle documents with different length. Third, in this study, we only used flow chart and Gantt chart for presentation in the ontotology-based system. There are other types of graphical charts that may better serve this purpose and should be explored in future research. Finally, this study only foscused on Chinese news. It would be necessary to validate the findings of this study using documents in other langauges.

Acknowledgements:

This project was partially funded by NSC 95-2416-H-110-001 in 2006, and Grant for research excellence in 2007-2009 to the first author from National Science Council. The authors also thank Mr. Yi-Cheng Yang, Bing-Huang Wu, and Kun-Han Chuang for their research assistance in the project.

References

Agrawal, R. and R. Srikant (1994). Fast algorithms for mining association rules. *20th VLDB Conference, Santiago, Chile, 1994*, pp. 487-499.

Antiqueira, L., O. N. Oliveira, L. F. Costa, and M. G. V. Nunes (2009). A complex network approach to text summarization Source. Information Sciences, (179)5, pp. 584-599.

Barzilay, R, Mckeown, K. and Elhadad, M. (1999). Information fusion in the context of multi-document summarization, in Proceedings of the 37th Annual Meeting of the Association for Computational Linguistics (ACL'99), College Park, Maryland, pp.550-557.

Blázquez, M., M. Fernández, J. M. García-Pinar, and A. Gómez-Pérez. (1998). Building Ontologies at the Knowledge Level using the Ontology Design Environment. *Knowledge Acquisition Workshop (KAW'98), 1998*.

Bloom, B. S. (ed.) (1969). *Taxonomy of Educational Objectives, the Classification of Educational Goals* Longman Group United Kingdom.

Brank, J., M. Grobelnik, and D. Mladenić (2005). A SURVEY OF ONTOLOGY EVALUATION TECHNIQUES. *Conference on Data Mining and Data Warehouses, Ljubljana, Slovenia, 2005*.

Brewster, C., H. Alani, S. Dasmahapatra, and Y. Wilks (2004). Data Driven Ontology Evaluation. *International Conference on Language Resources and Evaluation, Lisbon, Portugal, 2004*.

Brewster, C. and K. O'Hara (2004). Knowledge representation with ontologies: the present and future, *IEEE*

Intelligent Systems (19) 1, pp. 72 - 81.
Ceccaroni, L. and E. Kendall (2003). A graphical environment for ontology development. *The Second International Joint Conference on Autonomous Agents and Multiagent Systems*, Melbourne, Australia, 2003, pp. 958 - 959.
Davies, J., D. Fensel, and F. v. Harmelen (2002). *Towards the Semantic Web: Ontology-Driven Knowledge Management*. England: John Wiley & Sons Ltd.
Eom, J.-H. and B.-T. Zhang (2004) PubMiner: Machine Learning based Text Mining for Biomedical Information and Analysis, *Genomics and Informatics* (2) 2, pp. 99-106.
Felden, C. and D. Kilimann (2006). Deployment of Ontologies in Business Intelligence Systems. *The Eighth International Conference on Enterprise Information Systems, Paphos, Cyprus, 2006*.
Fortuna, B., D. Mladinic, and M. Grobelnik (2006) System for Semi-Automatic Ontology Construction. *The 3rd European Semantic Web Symposium,* Budva, Montenegro*, 2006*.
Goodman, K. S. (1973). Testing in reading: a general critique, in R. Ruddell (Ed.) *Accountability and Reading Instruction*: NCTE.
Gruber, T. R. (1993). A Translation Approach to Portable Ontology Specifications, *Knowledge Acquisition* (6) 2, pp. 199–221.
Guarino, N. and C. Welty (2004). An Overview of OntoClean, in S. Staab and R. Studer (Eds.) *Handbook on Ontologies*: Springer Verlag, pp. 151-159.
Hahn, U. and I. Mani (2000). The Challenges of Automatic Summarization, *Computer* (33) 11, pp. 29-36.
Hahn, U. and U. Reimer (1999). Knowledge-based Text Summarization: Salience and Generalization Operators for Knowledge Base Abstraction, in I. Mani and M. Maybury (Eds.) *Advances in Automatic Summarization*, Cambridge, Massachusetts: MIT Press, pp. 215-232.
Hartmann, J., Y. Sure, A. Giboin, D. Maynard et al. (2004). Methods for ontology evaluation. http://www.starlab.vub.ac.be/research/projects/knowledgeweb/KWeb-Del-1.2.3-Revised-v1.3.1.pdf
Hennig, L. W. Umbrath, and R. Wetzker (2008). An Ontology-Based Approach to Text Summarization. IEEE/WIC/ACM International Conference on Web Intelligence and Intelligent Agent Technology (WI-IAT '08)**.** Dec. 9-12, 2008. Sydney, NSW. 3, pp. 291-294.
Hevner, A., S. March, J. Park, and S. Ram (2004). Design Science Research in Information Systems, *Management Information Systems Quarterly* (28) 1, pp. 75-105.
Inniss, T. R., J. R. Lee, M. Light, M. A. Grassi et al. (2006). Towards applying text mining and natural language processing for biomedical ontology acquisition. *The 1st International Workshop on Text Mining in Bioinformatics,* Arlington, Virginia, USA, *2006*, pp. 7 - 14.
Karoui, L., M.-A. Aufaure, and N. Bennacer (2006). Context-based Hierarchical Clustering for the Ontology Learning. *The 2006 IEEE/WIC/ACM International Conference on Web Intelligence 2006*, pp. 420-427.
Lin, C. and Hovy, E.H. (2002). From Single to Multi document Summarization: A Prototype System and its Evaluation. In Proceedings of ACL-2002. Philadelphia, PA, USA. July 6-12, 2002.
Mani, I. and Bloedorn, E. (1999). Summarizing similarities and differences among related documents. Information Retrieval, (33)1, pp.1-23.
Mannila, H., Toivonen, H. and Verkamo, A.I. (1997). Discovery of Frequent Episodes in Event Sequences, *Data Mining and Knowledge Discovery*, 1, pp. 259-289.
McKeown, K. and Radev, D. (1995). Generating summaries of multiple news articles, in Proceedings of the 18th Annual International ACM SIGIR Conference on Research and Development in Information Retrieval (SIGIR 1995), ACM Press, New York, NY, pp. 74-82
McKeown, K., R. J. Passonneau, D. K. Elson, A. Nenkova (2005). Do summaries help? *Proceedings of the 28th annual international ACM SIGIR conference on Research and development in information retrieval (SIGIR' 05),* Salvador, Brazil, 2005, pp. 210-217.
Missikoff, M., R. Navigli, and P. Velardi (2002). Integrated approach to Web ontology learning and engineering, *IEEE Computer* (November), pp. 60-63.
Nastase, V. (2008). Topic-driven multi-document summarization with encyclopedic knowledge and spreadingactivation. In Proceedings of EMNLP 2008. Oct. 25-27, 2008. Waikiki, Honolulu, Hawaii.
Navigli, R., P. Velardi, and A. Gangemi (2003). Ontology Learning and Its Application to Automated Terminology Translation, *IEEE INTELLIGENT SYSTEMS,* (18) 1, pp. 22 - 31
Newell, A. (1982). The Knowledge Level, *Artificial Intelligence* (18) 1, pp. 87–127.
Radev, D., Jing, H., and Budzikowska, M. (2000). Centroid-based summarization of multiple documents: sentence extraction, utility-based evaluation and user studies, in Proceedings of ANLP/NAACL 2000 Workshop on Automatic Summarization, Seattle, WA, pp.21-29.

Saggion, H. and G. Lapalme (2002). Generating Indicative-Informative Summaries with SumUM, *Computational Linguistics: Special issue on Summarization* (28) 4, pp. 497-526.

Stevens, R., C. Goble, I. Horrocks, and S. Bechhofer (2002). Building a bioinformatics ontology using OIL, *IEEE Transactions on Information Technology in Biomedicine* (6) 2, pp. 135 - 141

Uschold, M. and M. King (1995). Towards a Methodology for Building Ontologies. *In Proceedings of Workshop on Basic Ontological Issues in Knowledge Sharing (IJCAI-95),* Canada.

Ustymenko, S. (2005). Ontology-supported sharing and reuse of knowledge in diverse environments. *ACM-SE 43 (43rd Annual Southeast Regional Conference),* Kennesaw, Georgia, 2005, pp. 347 - 348.

Vargas-Vera, M. and D. Celjuska (2004). Event Recognition on News Stories and Semi-Automatic Population of an Ontology. *2004 IEEE/WIC/ACM International Conference on Web Intelligence Beijing,* China, 2004, pp. 615 - 618.

Velardi, P., P. Fabriani, and M. Missikoff (2001). Using text processing techniques to automatically enrich a domain ontology. *2nd International Conference on Formal Ontology in Information Systems (FOIS 2001),* pp. 270-284.

Wayne, C. L. (2000). Topic detection and tracking in English and Chinese. *The Fifth International Workshop on Information Retrieval with Asian Languages,* Hong Kong, China, pp. 165 - 172.

Xu, M. (2002). Research on automated ontology construction based on semantic network, National Sun-Yet Sen University.

Yang, J., L. Wang, S. Zhang, X. Sui et al. (2004). Building Domain Ontology Based on Web Data and Generic Ontology. *IEEE/WIC/ACM International Conference on Web Intelligence,* Beijing, China, 2004, pp. 686 - 689.

Zhou, L. (2007). Ontology learning: state of the art and open issues, *Information Technology and Management,* (8) 3, pp. 241-252.

Appendix I. Screenshots of Two Systems Used in the Experiment

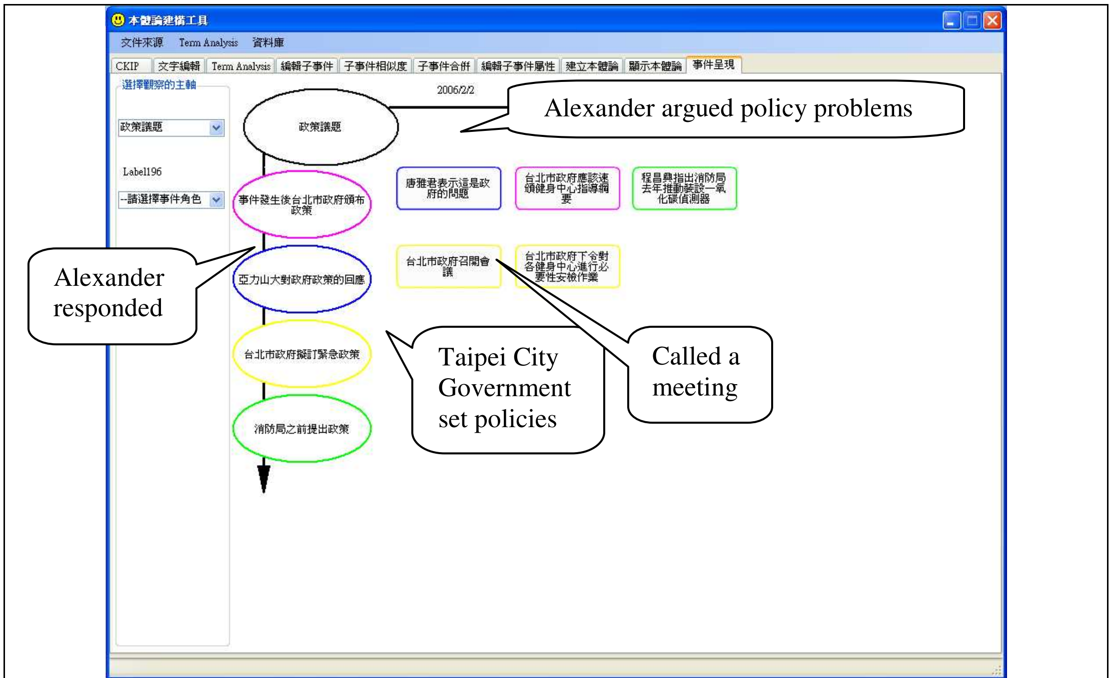

Figure 5. The Screenshot of Ontology System Interface

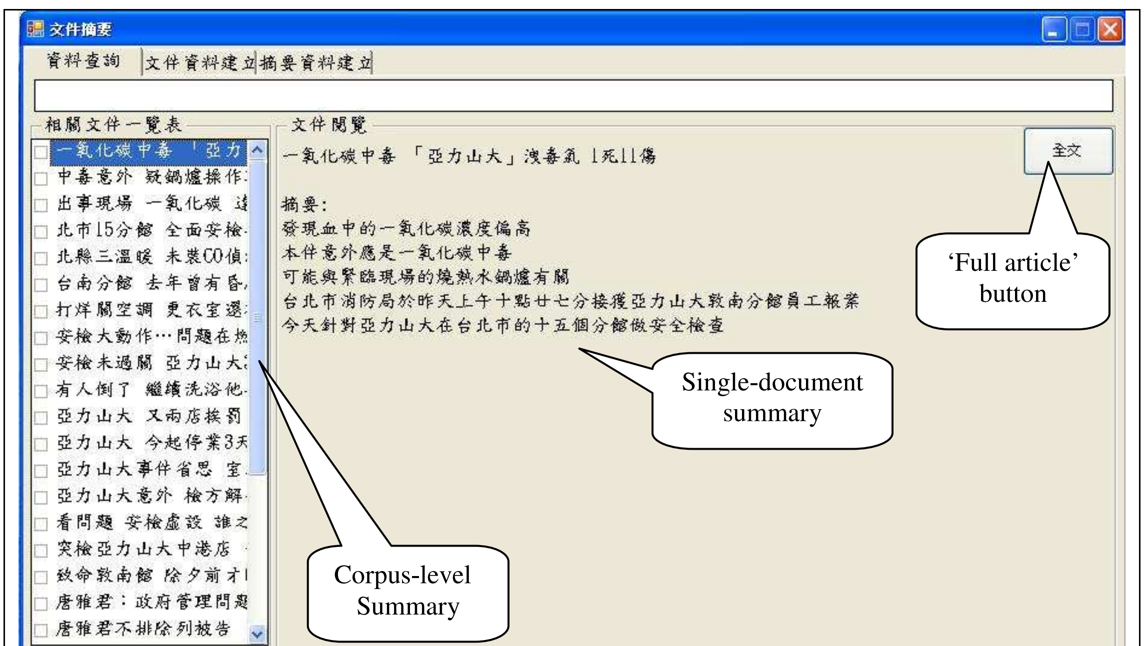

Figure 6. The Screenshot of Text Summarization System Interface

Appendix II: Question examples (Translated from original questions in Chinese)

Questions about factual knowledge:

1) In order to alleviate the carbon monoxide poisoning incidents, the Taipei Fire Department motivates residents to install carbon monoxide detector by offering ___ per household.

a) NT$ 300 b) NT$ 400 c) NT$ 500 d) NT$ 600 e) NT$ 700

2) How many people were died and injured in Alexander gas leaking accident?

a) One died, ten injured
b) Two died, eleven injured
c) Two died, ten injured
d) One died, eleven injured
e) One died, twelve injured

3) How many years will BenQ Corp. be authorized to use 'BenQ-Siemens' brand name after it acquires the Siemens Mobile Division?

a) 3 years b) 4 years c) 5 years d) 6 years e) 7 years

Questions about conceptual knowledge:

1) What are the most common accidents associated with boiler?

a) The poisoning gas leaking and fire
b) Incomplete burning and explosion
c) Fire and explosion
d) Incomplete burning and burn injuries
e) The poisoning gas leaking and burn injuries

2) Why doesn't BenQ have core communication technology?

a) Does not invest in R & D.
b) Competitors are too strong
c) Got in the market too late
d) Top management is not interested in communication market.
e) Does not have sufficient expertise in communication technology.

Questions about procedural knowledge (Participants were asked to produce the procedure of a process by ordering three activities):

1) After the gas leaking accident, what was the procedure for investigating the reason of the accident?

a) The Fire Department examined the gas.
b) Taipei mayor demanded public prosecutor to investigate the accident.
c) The police Department started collecting evidence on the boiler use and ventilation system.

2) What was the BenQ's procedure after its German cell phone factory claimed bankruptcy?

a) BenQ sold stocks of its subsidiary companies.
b) Requested bankruptcy protection.
c) Decided no longer responsible for paying off German company's deficit.