INFORMATION LEARNING WITH KNOWLEDGE MAPS

Jae-Hwa Lee
Department of Knowledge Service Engineering,
KAIST, Korea
tkod119@kaist.ac.kr

Aviv Segev
Department of Knowledge Service Engineering,
KAIST, Korea
aviv@kaist.edu

Keywords: knowledge maps, e-learning, knowledge representation.

ABSTRACT

Maps such as concept maps and knowledge maps are often used as learning materials. These maps have nodes as key concepts and links as relationships between key concepts. From a map, the user can recognize the important concepts and their interrelationships. To build concept or knowledge maps, domain experts are needed and the cost of map creation is high. This study developed a model which automatically builds a domain knowledge map (K-map) from a set of documents about a specific topic using text mining techniques. To build K-map, keywords are extracted from a set of documents using the TF/IDF algorithm and keyword pairs are ranked according to number of appearances in a sentence and number of words in a sentence. K-map was implemented to evaluate learning performance. The experiments analyzed both sentence recall and identification of important sentences, in comparison to document-based learning methods. The results show that K-map provides a mechanism with high recall and can distinguish the more important information, in comparison to document-based learning.

INTRODUCTION

When people learn from text, they usually follow the order set by the author, as with reading books, in the most common method of text-based learning. However, text-based learning is not efficient in the following situations. First, in many cases people have different levels of prior domain knowledge. However, when they learn from textual material, they can only read it from beginning to end or use the table of contents to jump to a specific chapter. If a learner has a certain knowledge level, he can use the index to look for information on a certain concept, but the index usually contains hundreds of concepts listed alphabetically with no relational information between them. Second, in cases of learning under time pressure the learner can use the table of contents or index to identify main parts. There is no other specific way to distinguish important information. Furthermore, if a person wants to learn about a domain from a web search, the time limitation is more acute due to the huge amount of documents online. He will read documents from top-ranked ones down and will stop when the time is up. Third, if a document is complicated or long, a reader may find it difficult to recognize important concepts and their interrelationships.

Map-based learning can be useful in these situations. Key concepts and relationships can be identified directly from the map. Time can also be saved by shortening the amount of text. However, these maps are built by the manual effort of domain experts. This paper proposes a method of automatic map generation and presents an implementation example with real-world data.

RELATED WORK

A concept map is a visuospatial representation of knowledge with text and graphical elements such as arrows, lines, ovals, and squares. It consists of nodes, containing a concept or item, and links connecting two nodes to each other and describing their relationship, where each node-link relation makes a proposition.

A concept map is taken from the theories of assimilation and subsumption in cognitive learning theory (Ausubel, 1968). A concept mapping tool was initially developed to search for better ways to represent learner knowledge (Novak & Musonda, 1991). A topic map is an ISO standard for describing knowledge structures and associating them with their resources and contains basic concepts, such as Topics, Associations, and Occurrences (Pepper, 2000). An environment for e-learning, called TM4L, where people use topic maps for learning, was developed in (Dicheva & Dichev, 2006).

A way to construct concept maps automatically from academic papers was proposed in (Chen et al., 2008). They used author keywords as keywords after pre-processing and defined relations among them with four assumptions. While previous work dealt with constructing maps or using maps for learning, this study focuses on automatic map construction for learning.

MODEL OVERVIEW

K-map is defined as a domain knowledge description map with nodes and links. Nodes are keywords considered important domain concepts and links are relations between two keywords. The present work developed K-map Tools, a learning environment where the user controls the number of keywords and relations, and thus the requested depth of information, that are used to automatically produce a K-map for a document. There is a concept search window that helps the user search for a specific concept. The system takes several documents related to a domain and generates a K-map. Figure 1 is an example of a K-map produced by K-map Tools, with 18 keywords and 30 relations, generated from a set of documents about John F. Kennedy. Each link has a different score and thickness, when a higher score yields a thicker link.

The K-map has a hyperlink on each relation. If the user clicks a relation, he will see all sentences having the two keywords at both ends of the relation. For example, if a relation between 'Kennedy' and 'president' is chosen, Figure 2 appears. In Figure 2, all sentences having the words 'Kennedy' and 'president' are shown. The user can read some of the sentences and acquire domain knowledge.

With K-map Tools, as the user handles the K-map, he searches and learns. In the searching and learning process with current search engines, the user first types in words he wants to know about; these can be considered the topic. Then the search engine shows him a list of documents related to the query. The user goes over the list, at some point choosing a document and reading. He repeats the process until he learns enough. With current search engines, there may be inefficiencies – the user may not understand important concepts, may miss key concepts, or may waste his time trying to find a right document.

However, with K-map the user doesn't have to find a right document, because the map already has extracted sentences, categorized by keyword pairs. He can recognize the key concepts and the strongly connected key concepts; he obtains a holistic view of the domain. Thus, K-map improves the searching and learning process.

IMPLEMENTATION AND EXPERIMENTS

The K-map model was implemented as a system. The model consists of three parts: keyword extraction, relation extraction, and relation labeling. For the experiments 34 participants were divided into two groups: a document group which learnt from a set of documents and a map group which learnt from K-map using K-map Tools.

Three experiments were performed. The first experiment analyzed whether the K-map could identify the important sentences in a document. The results showed that the model can successfully filter out the sentences considered not important to the main idea. The second experiment analyzed in the two participant groups the recall - the number of important sentences identified. The results show that there was no statistical difference between the groups. The third experiment compared the number of significant sentences recalled between the two groups. The results showed that with K-map users learned information that is more important, in comparison to the information learned with documents.

The results showed that, contrary to the initial assumption, with K-map people do not read faster or recall more information compared to people who learn with texts. It is assumed that the learning speed using K-map depends on previous experience and might increase as users become more experienced. Further studies can investigate the learning curve of using K-map and analyze K-map learning performance in different domains.

DISCUSSION AND CONCLUSION

K-map offers multiple benefits. A user can see key concepts in a domain as well as strongly related concepts. As a user reads, he can directly access a document he wants from a certain sentence; in other words, K-map can function as a search engine. By exploring the map, a user can learn about the domain at some level of knowledge without accessing original documents. As a user explores a domain K-map, he can see the holistic/overall picture. Additionally, since a user can choose relations based on keywords, he can selectively learn about the domain, which is hardly possible when learning from text. Furthermore, experiments show that K-map provides a mechanism with which to distinguish the more important sentences.

REFERENCES

Ausubel, D. (1968). *Educational psychology: A cognitive approach*. New York: Holt, Rinehart and Winston.

Chen, N., Kinshuk, Wei, C., Chen, H. (2008). Mining e-Learning domain concept map from academic articles, *Computers & Education*, 50(3), 1009-1021.

Dicheva, D., Dichev, C., (2006). TM4L: Creating and browsing educational topic maps, *British Journal of Educational Technology*, 37(3), 391-404.

Novak, J., Musonda, D. (1991). A twelve-year longitudinal study of science concept learning. *American Educational Research Journal*, 28, 117-153.

Pepper, S., (2000). The TAO of topic maps. *Proceedings of XML Europe*.

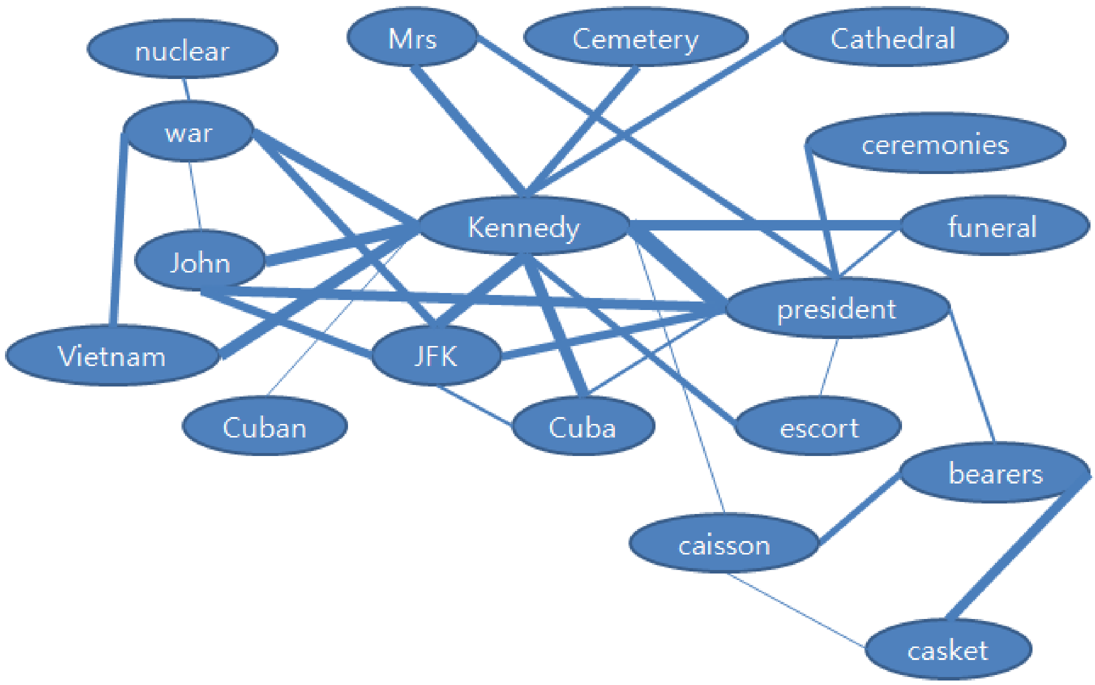

Figure 1. Knowledge map of John F. Kennedy

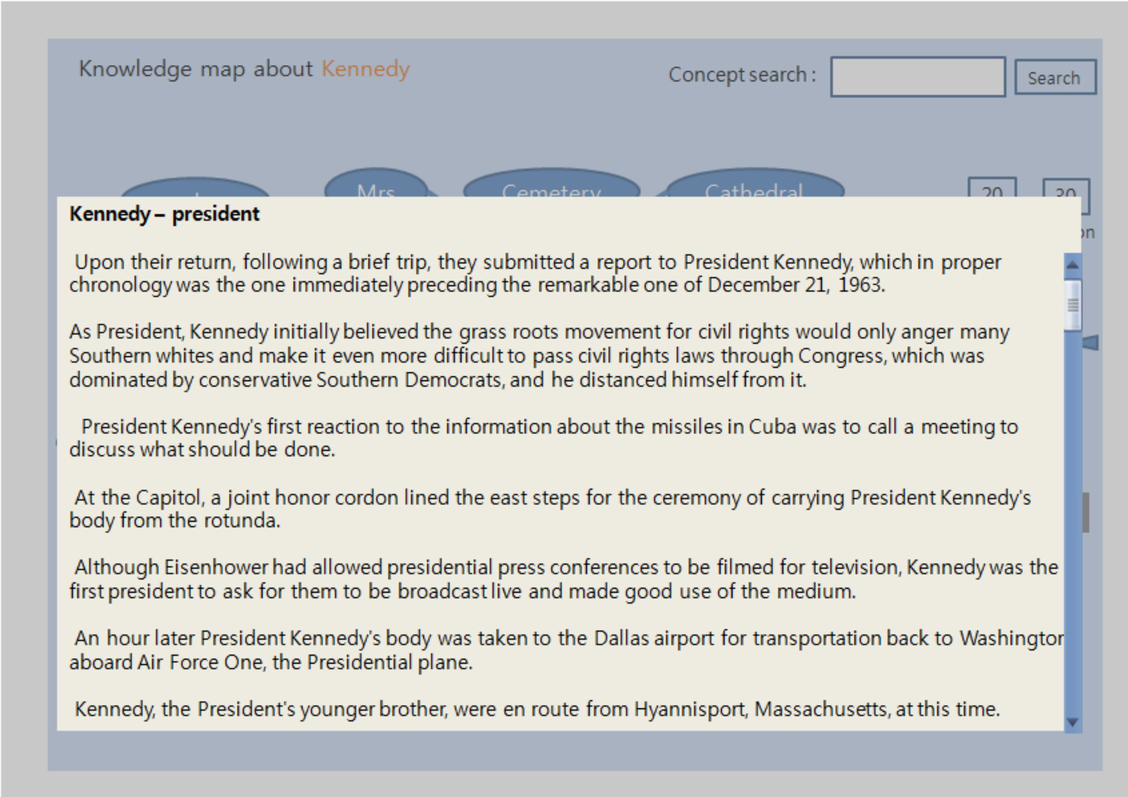

Figure 2. Screen shot of sentences containing 'Kennedy' and 'president' in K-map Tools