Computers & Education 59 (2012) 353–364

Contents lists available at SciVerse ScienceDirect

Computers & Education

journal homepage: www.elsevier.com/locate/compedu

Knowledge maps for e-learning

Jae Hwa Lee, Aviv Segev*

Department of Knowledge Service Engineering, KAIST, Republic of Korea

ARTICLE INFO

Article history:
Received 16 August 2011
Received in revised form
23 January 2012
Accepted 25 January 2012

Keywords:
Knowledge maps
E-learning
Concept relation

ABSTRACT

Maps such as concept maps and knowledge maps are often used as learning materials. These maps have nodes and links, nodes as key concepts and links as relationships between key concepts. From a map, the user can recognize the important concepts and the relationships between them. To build concept or knowledge maps, domain experts are needed. Therefore, since these experts are hard to obtain, the cost of map creation is high. In this study, an attempt was made to automatically build a domain knowledge map for e-learning using text mining techniques. From a set of documents about a specific topic, keywords are extracted using the TF/IDF algorithm. A domain knowledge map (K-map) is based on ranking pairs of keywords according to the number of appearances in a sentence and the number of words in a sentence. The experiments analyzed the number of relations required to identify the important ideas in the text. In addition, the experiments compared K-map learning to document learning and found that K-map identifies the more important ideas.

1. Introduction

When people learn from textual material, they usually follow the order set by the author, as with reading books. Although this is the most common method of text-based learning, it is not efficient in the following three situations. First, in many cases, people have different levels of prior domain knowledge. However, when they learn from text material, they can only read it from the beginning to the end or use the table of contents to jump directly to a specific chapter. If a learner has a certain level of knowledge, then the index can be used to look for the information on a certain concept. However, the index usually contains hundreds of concepts listed in alphabetical order 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 the 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, then the time limitation is more critical due to the huge number of documents on the web. Learners can read documents from top-ranked ones down and then stop reading when the time is up. Third, if a document is complex or long, then a reader may find it difficult to recognize the important concepts and the relationships among them.

Concept maps or knowledge maps can be useful in these situations and thus can improve the e-learning experience. Key concepts and relationships can be recognized directly from the map, so learners can identify them with minimum effort. Time can also be saved when the amount of text is shortened.

However, the construction of concept maps and knowledge maps requires manual effort of domain experts. In this paper, a method for automatic generation of maps is proposed and an example of implementation with real-world data is presented.

2. Related work

The section presents the different types of maps currently used for knowledge representation: Concept Map, Knowledge Map, and Topic Map. The advantages and limitations of using each map type are discussed. The last section presents previous work on automatic construction of maps relevant to the field of learning.

* Corresponding author.
E-mail address: aviv@kaist.edu (A. Segev).

0360-1315/$ – see front matter
doi:10.1016/j.compedu.2012.01.017

2.1. Concept map

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.

Concept maps are theoretically grounded in cognitive learning theory. Based on Ausubel's theories of assimilation and subsumption (Ausubel, 1968), which state that the most important single factor influencing learning is what the learner already knows, Novak and Musonda (1991) initially developed concept mapping tools to search for better ways to represent the learners' knowledge.

Since then, concept maps have been used in various areas such as assessment tools (McClure, Sonak, & Suen, 1999), cooperative learning tools (Stoyanova & Kommers, 2002), anxiety reduction tools (Czerniak, 1998), and tools that increase efficiency of search engines (Carvalho, Hewett, & Cañas, 2001). Computer-based environments for concept mapping have been developed, such as the dynamic concept map proposed by Nesbit and Adesope (2005), in which audio presentation is synchronized with a node-link image. Cañas et al. (2005) developed CmapTools, tools that enable concept maps to be combined with multimedia resources. Chu, Lee, and Tsai (2011) developed a concept map learning system with ontology technology to help users search the concept map to help reduce the user's cognitive load. Ruiz-Martínez, Valencia-García, Fernández-Breis, García-Sánchez, and Martínez-Béjar (2011) developed the MCRDR tool for concept acquisition in the biomedical domain. However, these methods did not automatically create the concept map for learning.

2.2. Knowledge map

Holley and Dansereau (1984) first developed knowledge maps, and Wiegmann, Dansereau, McCagg, Rewey, and Pitre (1992) added types of links, which include type, characteristic, part, results in, leads to, and example. Knowledge maps are very similar to concept maps, in terms of structures, goals, benefits, and areas of application; however, knowledge maps have specific types of links. A knowledge map is also a two-dimensional graphical display that presents information, but the method of relationship designation in knowledge maps distinguishes the technique from other types of maps such as concept maps (Amer, 1994). O'Donnell, Dansereau, and Hall (2002) summarized findings on the field of knowledge maps over 12 years: students who use knowledge maps recall a greater number of central ideas than do students who use texts (Hall, Dansereau, & Skaggs, 1992); students with low verbal ability or low prior knowledge benefit the most from the presentation of information in a knowledge map format with recall (Rewey, Dansereau, Skaggs, Hall, & Pitre, 1989); students who use knowledge maps as supports when interacting with peers in cooperative learning environments learn more effectively; and information is recalled better when presented in well-structured maps designed according to Gestalt principles than when presented in less well-structured maps (Wiegmann et al., 1992). Also, knowledge maps were found to help general purpose conceptualization processes (Gomez, Morenoa, Pazosa, & Sierra-Alonso, 2000) and to help users reduce their anxiety and increase motivation (Hall & O'Donnell, 1996).

2.3. Topic map

A topic map is an ISO standard (ISO/IEC 13250, 2003) for describing knowledge structures and associating them with their resources. A topic map contains basic concepts, such as Topics, Associations, and Occurrences (Pepper, 2000). A topic map also consists of topics and relationships between them. If the word 'topic' is generalized to 'word', topic maps are more or less similar to concept maps, except that topic maps are more focused on standards. Böhm, Heyer, Quasthoff, and Wolff (2002) introduced a way to construct an initial set of topic maps or extend/optimize a given map using text mining technology. Dicheva and Dichev (2006) developed an environment for e-learning, called TM4L, where people use topic maps for learning. The TM4L is an environment for building, maintaining, and using standards-based, ontology-aware e-learning repositories, based on the idea that concept-driven access to learning material implemented as a Topic Map can bridge the gap between a learner and targeted knowledge. The goal was to support an efficient context-based retrieval of learning content tailored to the needs of a learner working on an educational task. Other learning environments include a ubiquitous learning system (Wang & Wu, 2011), English e-learning system (Wang & Liao, 2011), and performance analysis e-learning system (Jia et al., 2011).

2.4. Automatic construction of maps

In recent years interest in automatically building concept maps has grown. Chen and Xia (2009) reviewed research about automatic construction of concept maps. They first presented a traditional concept map construction method achieved by hand which includes the following steps: concept selection, concept classification, appointing the central concept, and connecting concepts and cross-concepts, and then described auto-construction concept map methods based on document information retrieval and concept extraction. Chen, Kinshuk, Wei, and Chen (2008) proposed a way to construct concept maps automatically from academic papers. They used author keywords as keywords after pre-processing and defined relations among them with four assumptions:

Each keyword listed in a research article represents one essential concept.

If two keywords appear in one research article, it implies that a certain relation exists between these two keywords.

The higher the frequency of occurrences of two keywords appearing in one sentence, the higher the relation is between them.

The shorter the 'distance' between two keywords in one sentence, the higher the relation is between them.

To define a relation in this study, two factors are considered, implicitly assuming the first two assumptions. One factor is how frequently two keywords appear in a sentence together, which is the same as the third assumption. The other factor is how big a role the two keywords play in a sentence. As the number of the words in a sentence increases, the weight of each word decreases. In other words, the score of a relation in a shorter sentence is higher than the score of the relation in longer sentences. We excluded the fourth assumption, because our experiments showed no direct relation between the 'distance' of words in a sentence and their semantic relation. Each relation has its own score, and it is used to rank the relations.

Tseng, Sue, Su, Weng, and Tsai (2007) proposed a Two-Phase Concept Map Construction (TP-CMC) approach to automatically construct a concept map for a course from the learner's past test data. The first phase uses Fuzzy Set Theory to transform the numeric data into symbolic ones. The second phase uses multiple rule types to analyze the mined association rules and also uses a heuristic algorithm to remove redundancy and circularity when building concept maps. Relations in the concept map indicate learning paths. Hou, Ong, Nee, Zhang, and Liu (2011) proposed GRAONTO, generating graphs of documents and using random walk term weighting to estimate the relevance of the information of a term to the corpus from both local and global perspectives. The Markov Clustering algorithm was used to disambiguate terms with different meanings and group similar terms to produce concepts. Urfat and Korhan (2009) introduced a way to extract key concepts and relations among learning concepts. To extract key concepts, n-gram, a sequence of characters that stands for a word, was used with a dictionary of technical terms. To extract relations (relevance) from sentences, three features were considered: word co-occurrences, logarithmic concept weights, and augmented normalized candidate learning concept frequency.

Previous work focused on automatic construction of maps and on construction of maps for learning based on user past learning experience. This paper proposes a method of automatic map generation from unexplored material for e-learning purposes and presents an implementation example with real-world data.

3. Model

3.1. Overview

K-map is a knowledge map that has nodes and links. Nodes are keywords that are considered important concepts for a specific domain, and links are relations between two keywords. The system uses several documents related to a certain domain to generate a K-map. Fig. 1 is an example of the K-map Tools screen shot automatically generated from a set of documents about John F. Kennedy. The map has 18 keywords and 30 relations. Links have different thicknesses. Each link has a different score and thickness, when a higher score yields a thicker link. K-map Tools serves as a K-map learning environment.

K-map has a hyperlink for each relation. If the user clicks a relation, all the sentences that have two keywords at both ends of the relation will be visible. For example, if a relation between 'Kennedy' and 'president' is chosen, Fig. 2 appears. In Fig. 2, all sentences that have the words 'Kennedy' and 'president' are displayed. The user can read some of the sentences and acquire domain knowledge from them. If the user chooses any sentence, the sentence can be directly accessed in the original document (Fig. 3).

With K-map Tools, the user can control the number of keywords and relations. Keywords and relations will appear and disappear according to their ranking and the limit that the user sets. This function assists in reducing map shock (Dansereau, Dees, & Simpson, 1994) since the user can resize the map freely. There is a concept search window that helps the user search for a specific concept. When a relation has too many sentences, a user can be overwhelmed and lose interest. In this case, the words extracted from the set of sentences are believed to represent the relation. There are five extracted words for the relationship between 'Kennedy' and 'Vietnam' (Fig. 4). If the user chooses one of those words, K-map Tools will show the sentences that have 'Kennedy', 'Vietnam', and the word chosen. In Fig. 5, the word is 'withdraw'. These words categorize the initial sentences.

With K-map Tools, as the user handles the K-map two tasks are accomplished: searching and learning. In current search engines two steps occur in the searching and learning process. The first step is the user typing in the words the user wants to know about; those words can be considered the topic. Then the search engine shows the user a list of the documents related to the query. The user goes over the list

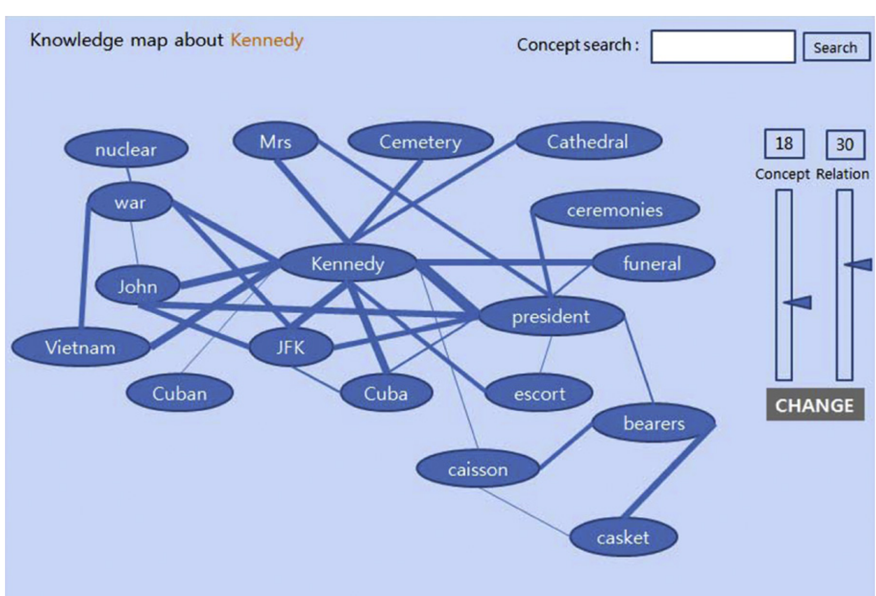

Fig. 1. Screen shot of K-map tools - knowledge map of John F. Kennedy.

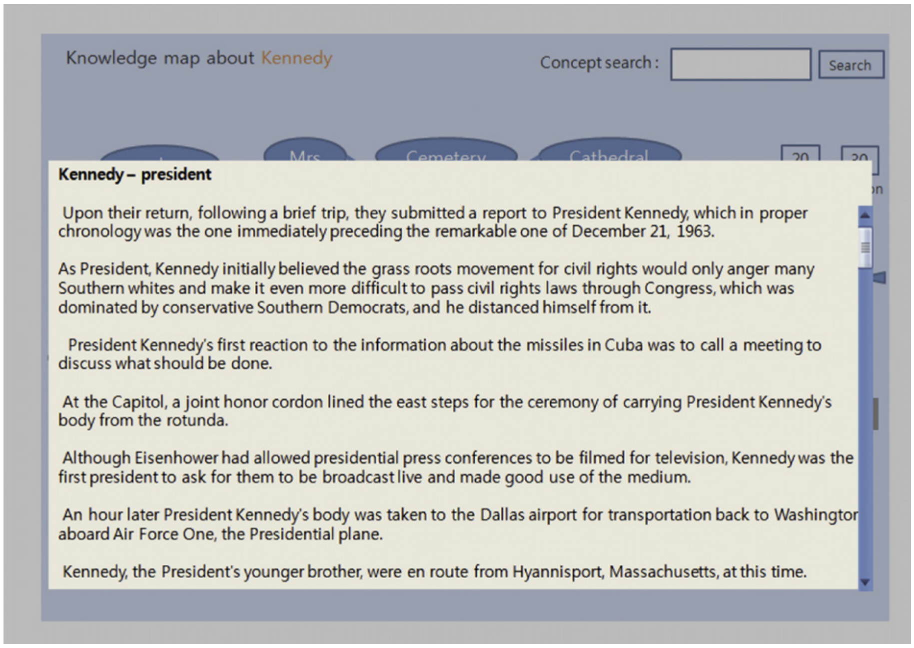

Fig. 2. Screen shot of sentences containing 'Kennedy' and 'president' in K-map tools.

and at some point chooses a document and starts reading it. The same process is repeated until enough has been learned about the topic. With current search engines, there may be inefficiencies – the user may not understand important concepts, may miss key concepts, or may waste 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. The user can recognize the key concepts and the strongly connected key concepts and obtain a holistic view of the domain. Thus, K-map improves the searching and learning process.

The sentences are organized according to key concepts. The user selects the key concepts to learn about and then reads the sentences that are relevant to these concepts. Therefore, in most cases the context of the sentences is already defined and presented to the user through the use of the concepts and relations. After the user sees visually the context of the sentences as defined by the concepts and relations, the user can then read them.

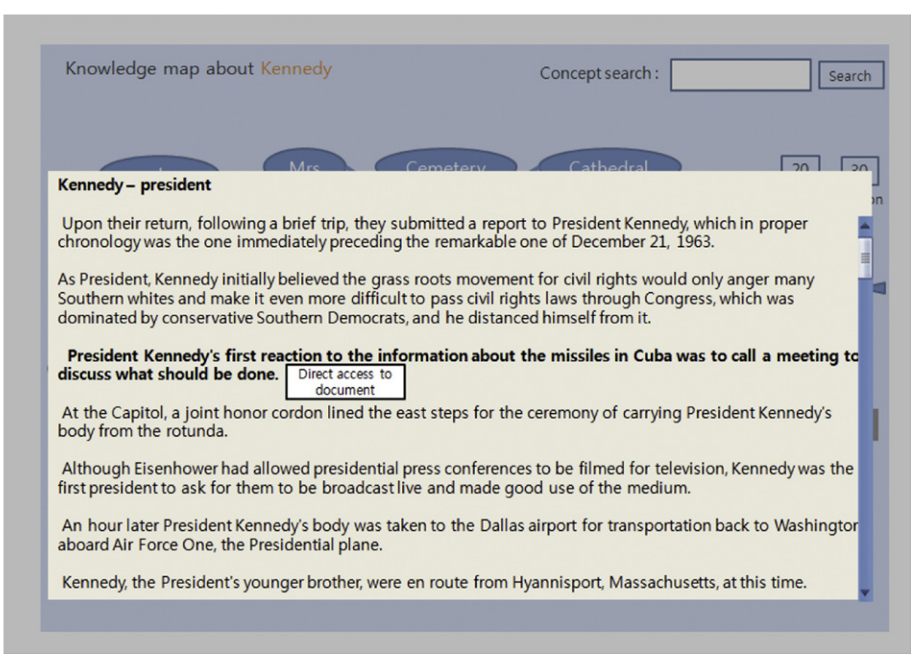

Fig. 3. Screen shot of K-map tool direct access to a document through a sentence.

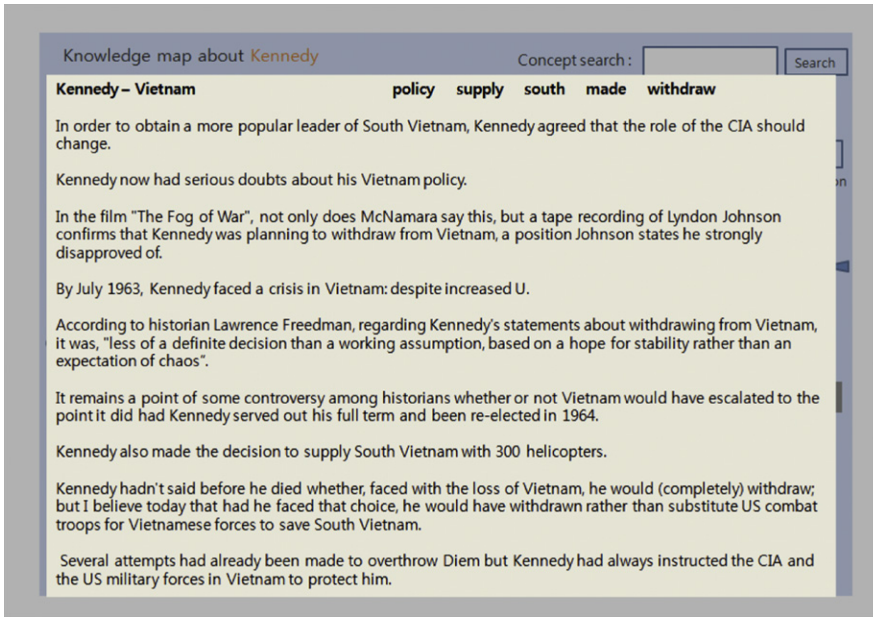

Fig. 4. Screen shot of K-map tools categorizing sentences with more words.

3.2. Model structure

In this section, the processes of K-map construction are described. Model construction consists of three parts, which are keyword extraction, relation extraction, and relation labeling.

3.2.1. Keyword extraction

In the first phase, keywords are extracted from a set of documents. The documents can be manually chosen or chosen from the user's query. Each term will receive a weight using the term weighting algorithm. Keywords will be selected from the top-ranked terms set by a limit given by the K-map user. Before running the term weighting algorithm, stop words, such as "and" and "the", are deleted from original sentences of a document. Next, a stemming process is conducted with those sentences that don't have stop words. The stemming process cuts the original word down to the root. For example, after the stemming process, the words "run", "running", and "runs" become the same word, which is "run", because all the words have the same meaning. To stem words, Porter's algorithm was used in the K-map system (Porter, 1980). Once these pre-processes are finished, we can use a term weighting algorithm to weight terms. The weighting algorithm is based on the Term Frequency/Inverse Document Frequency (TF/IDF) method (Salton & Buckley, 1988). The formula below is the weighting method using TF/IDF.

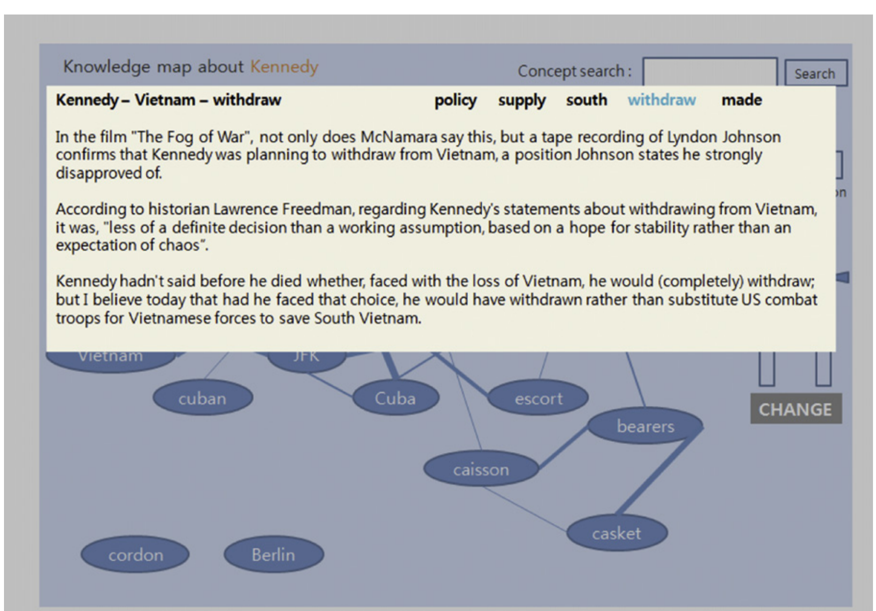

Fig. 5. Screen shot of K-map tools categorizing sentences with added word 'withdraw'.

$$w_{ik} = \frac{tf_{ik}\log(N/n_k)}{\sqrt{\sum_{k=1}^{t}(tf_{ik})^2[\log(N/n_k)]^2}}$$

W_{ik}: weight of term k in document i
tf_{ik}: term frequency of term k in document i
N: total number of documents
n_k: number of documents that contains term k

After the weighting process, each term has its own weight. A similar word can have different weights in different documents due to its frequency in a specific document, since the word is weighted for each document using the TF/IDF algorithm. If the same word appears in multiple documents, its weight is determined by the maximum value for all of its weight.

$$W_{MT} = Max(W_{D_iT})$$

$\mathrm{D}_i = i$-th document, $i = 0,1,2\ldots$total number of the documents in K-map
$W_{\mathrm{T}} =$ weight of term T in K-map
$W_{D_iT} =$ weight of term T in D_i

Once words are ordered by weights, keywords are selected from top-ranked words according to the limit that the user set. Unlike concept and knowledge maps, K-maps can have other parts of speech, not only nouns, as their nodes. We decided to take other parts of speech as well, because we thought that verbs also can have meaning if they are accorded a high weight.

3.2.2. Relation extraction

Once the keywords of the K-map are decided, relations are defined. Chen et al. (2008) identified assumptions for defining relations. In the present study, two factors are considered based on these assumptions. One factor is how frequently two keywords appear in a sentence together, and the other is how big a role the two keywords play in a sentence. As the number of the words in a sentence increases, the weight decreases. In other words, the score of a relation in a shorter sentence is higher than the score of a relation in longer sentences. Each relation has its own score, and it is used to rank the relations.

The characteristics of important sentences are generally that they contain more important words (or keywords) and have shorter sentence length and a lower number of sub-sentences (Li & Choi, 1997). Clark and Wasow (1998) investigated the origins of repeated words in spontaneous speech. They compared content words and function words. Content words refer to entities, events, states, relations, and properties in the world. They are characteristically nouns, verbs, adjectives, or adverbs. Function words, in contrast, are used largely to express the relations among elements of sentences or to indicate their discourse functions. They include prepositions, conjunctions, articles, auxiliary verbs, and pronouns. The results showed that only 0.0252% of the function words and only 0.0024% of the content words are repeated. It can be assumed that in writing the number of repeated words is even lower. These numbers of repeated words in a sentence seem to be low enough as to not have any major impact on the algorithm performance.

$$R_{i,j} = \sum_{D_m}\sum_{S_n}\frac{2}{N_{D_mS_n}}$$

$i, j =$ keyword pair
$R_{i,j} =$ score of relation between word i and word j
$m = 1,2,\ldots$ total number of documents in a map
$n = 1,2,\ldots$ total number of sentences in document Dm
$S_n =$ n-th sentence
$D_m =$ m-th document
$N_{\mathrm{D_mS_n}} =$ total number of words in sentence S_n, document D_m

3.2.3. Relation labeling

When the user chooses thicker lines, which are considered stronger connections, the user can access many sentences at the same time. This can be overwhelming for the user, and therefore a way of decreasing the number of sentences is suggested: to extract words representing the relation and to use them in categorization. Initially, several sentences were extracted. Within the sentences, for each term its term frequency is calculated. TF/IDF was empirically found to be less reliable than TF. The most frequent words, not including the two keywords, can be considered representative words. Once the user chooses one of these representative words, the user will see the sentences that have two keywords and all representative words. Representative words categorize the initial sentences. Categories might overlap one another. Resolving the category overlap requires further research.

4. Experiments

Two sets of experiments were performed to examine the K-map model. The first set of experiments examines whether the K-map model extracts the more important sentences in a text by analyzing the number of important sentences extracted versus the less important ones. The second set of experiments compares the results of learning performance between two groups, which are a Document-based learning group and a Map-based learning group.

4.1. Recall of important sentences

4.1.1. Method

In this experiment, the recall of important sentences of the K-map model was compared to human classification. Each sentence in the text is assigned to one of three classes (A, B, C) according to the importance of its information. Three graders were asked to classify the sentences. All the graders were graduate students with English as a second language and the same mother tongue. To decrease misunderstandings and biases, they were given enough time to read documents thoroughly and discuss the contents on the level of each sentence. In case of disagreement the decision was made based on the majority classification.

Each sentence was assigned its classification according to one of the following:

Class A: Sentences that have main ideas or play a major role in understanding the topic.
Class B: Sentences that support main ideas or are partially helpful in understanding the topic.
Class C: Sentences that are not related to the topic or are not helpful.

The experiments were conducted on two topics, a New York Times news article titled "Are Bad Knees in Our Genes?" (September 29, 2010, 986 words, 44 sentences) and an article from Wikipedia titled "Stock Market Bubble" (1044 words, 45 sentences).

4.1.2. Recall results of important sentences

4.1.2.1. News article. The article has a total of 44 sentences. After the three graders decided the class of the sentences, the article had 8 sentences of class A, 25 sentences of class B, and 11 sentences of class C. Then we examined two cases. The first case was using K-map with 8 concepts (3% of the total number of words in the article). The second case was using K-map with 24 concepts (9% of total words). Since relationships in K-map have rankings, the assumption was that the learner would learn from the first ranked relationship to the lower ones.

Fig. 6 shows the accumulative number of sentences that a user of K-map can reach, with the ranking information, when K-map has 8 concepts. *X*-axis represents the number of relations that a user accesses in the order of ranking. *Y*-axis is the number of sentences that a user accesses by choosing relations. The graph shows that sentences in class B are extracted more than others.

Fig. 7 represents the percent of important sentences extracted for each category. *Y*-axis is percentage of extraction for each level. Sentences in class A have the highest extraction rate, followed by sentences in class B. Sentences in class C have a far lower extraction rate.

The same experiment was conducted with 24-concept K-map. The results are displayed in Figs. 8 and 9. For example, if a user reads sentences from the top three relations, the user is going to access 22 sentences. Among them, the number of class B sentences is 15, the number of class A sentences is 6, and the number of class C sentences is only 1. In percentages, the sentence extraction percentage of class B is 60% and that of class A is 75%, while that of class C is 9.09%. The results emphasize the gap between the extraction rate of class C sentences and that of class A and B sentences.

4.1.2.2. Wikipedia article. The article has 45 sentences in total. After the three graders decided on the class of the sentences, the article had 14 sentences of class A, 20 sentences of class B, and 11 sentences of class C. Then we examined two cases. Fig. 10 presents the number of sentences that the user of 3% K-map (9 concepts) can reach and Fig. 11 the percentage of sentences using 3% K-map.

The results show that the extraction percentage is not that high, even with sentences in classes A and B. Since the map was built with 9 concepts, we can infer that this article has a greater number of important concepts. Similarly, Figs. 12 and 13 display the results with a map of 27 concepts (9% of total words). This time, overall extraction rate has increased as number of accessed relations increases. With the 33rd relation, extraction rate of class A sentences reached 100%, which means the user accessed all the main idea sentences, by accessing the top 33 relations.

4.2. Learning performance

4.2.1. Method

The second experiment examined learning performance according to free recall of main ideas. The number of extracted idea units from the Document group was compared to the number of extracted idea units from the Map group. The number of extracted not-important idea units was also compared.

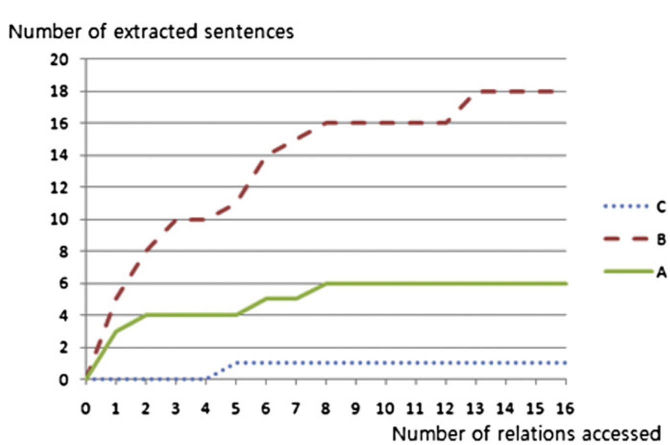

Fig. 6. Number of sentences extracted in 3% K-map.

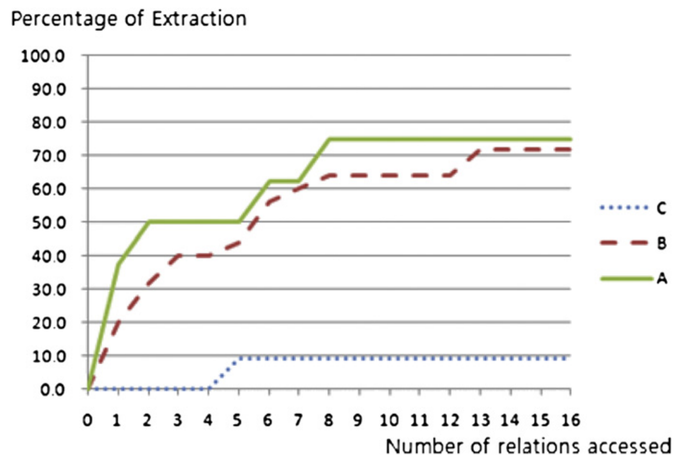

Fig. 7. Percentage of sentences extracted in 3% K-map.

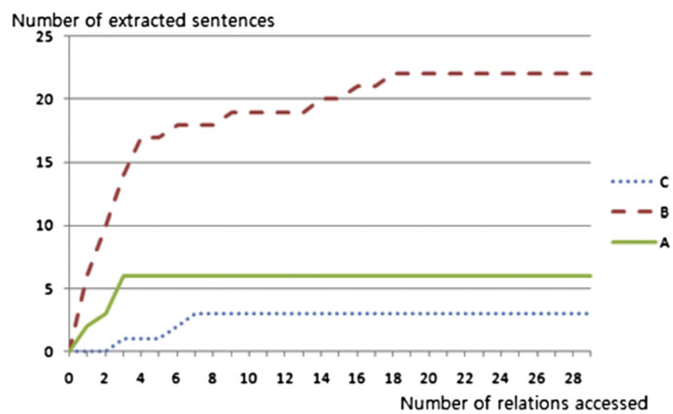

Fig. 8. Number of sentences extracted in 9% K-map.

To gather participants, on-line and off-line advertisements were posted at several places in a university campus one week before the experiment date. Only undergraduate and graduate students with English as a second language and the same mother tongue were accepted. Participants were randomly assigned to two groups. Furthermore, participants were allowed to use electronic dictionaries. A total of 35 students participated in the experiment. One participant was later found to have lived in an English-speaking country for more than 8 years, so this participant's result was excluded because it could influence the overall result. Except for this participant, one year of life in an English-speaking country was the longest time among participants and was not thought to have a critical effect that could distort the results.

Participants were asked to write everything they learned from the material after 8 min of learning time. Graders who did not have any information about the groups determined the free recall scores for all the participants. Since this experiment deals with free recall a different grading method was selected, 0–1–2, since these sentences were created by the participants and can include more than one important topic

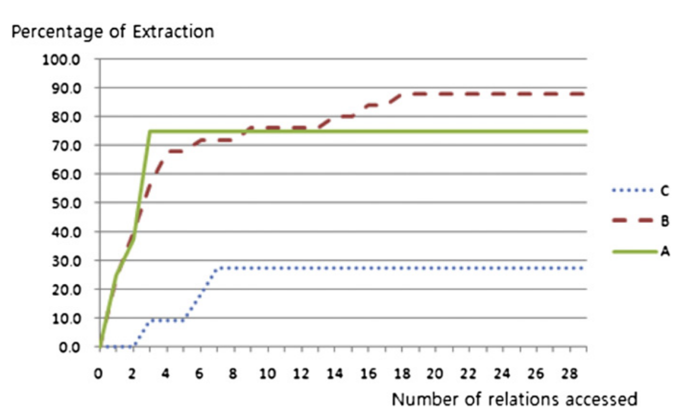

Fig. 9. Percentage of sentences extracted in 9% K-map.

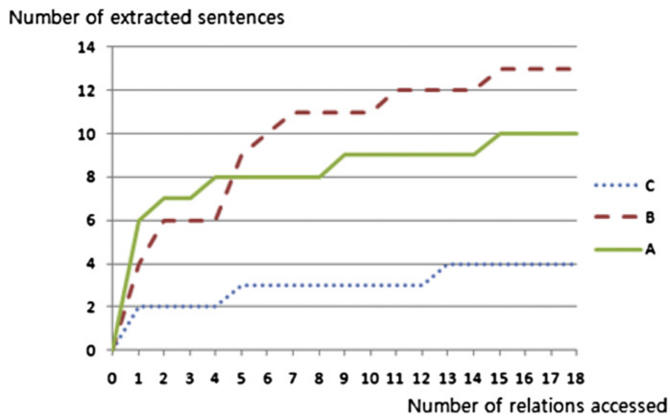

Fig. 10. Number of sentences extracted in 3% K-map.

and cannot be mapped directly to the A–B–C classification. The scoring technique is a modification of the scoring technique used in Dansereau et al. (1994). The graders read the participants' answer sheets sentence by sentence and divided each sentence into smaller ones so that each sentence had only one piece of information, which is called an *idea unit*. Then the graders matched each idea unit to the contents of the document set. The idea unit received a score ranging from 0 to 2, where:

0: The idea unit is completely inaccurate or cannot be matched to one of the idea units from the documents.
1: The idea unit is partially matched or partially accurate.
2: The idea unit is completely accurate and matched.

The total score of an answer sheet was the sum of the scores of the idea units.

4.2.2. Result of free recall test

Table 1 presents the results of the free recall test of both groups. Observed mean, standard deviation, and sample size are shown.

To statistically compare the results, Anderson–Darling test was conducted for normality. Both groups were found to have normality (Document group: $p = 0.208$, Map group: $p = 0.608$). Also, these two groups were found to have equal variances with F test ($p = 0.874$) and Levene test ($p = 0.955$). Finally, t-test was conducted with the two groups. The result indicates that there is no statistical difference in average score between the two groups ($p = 0.886$).

Table 2 analyzes the number of idea units between the two groups without considering how accurate each idea unit is.

The Map group and Document group both have normality (Document group: $p = 0.190$, Map group: $p = 0.695$). Also the two groups have the equal variance with F test ($p = 0.806$) and Levene test ($p = 0.954$). The result of t-test shows that there is no difference between the two groups ($p = 0.901$). In other words, there is no difference in the number of idea units extracted.

Initially, we expected that the recall in the Document group would be higher than the recall in the Map group. When the user learns with documents, the user follows the author's lead: in other words, the user reads a document from the beginning to the end. If the user learns with K-map, however, even though the user reads several sentences together on the same page, the user may have to put more effort into reading them, because there is no flow between sentences and each sentence is independent from others.

However, the result of this experiment did not indicate a difference in average score. Since the participants' background was similar, a possible reason can be the short period of time: since the experiment time was relatively short (8 min), participants did not lose concentration. If the time period is longer, then it is more likely that the score of the Map group would be higher than that of Document

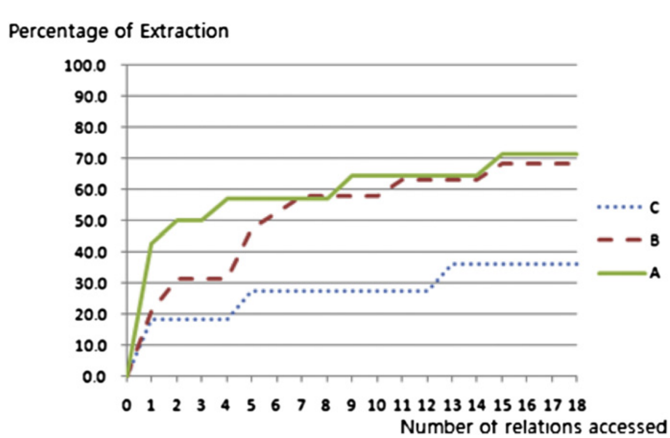

Fig. 11. Percentage of sentences extracted in 3% K-map.

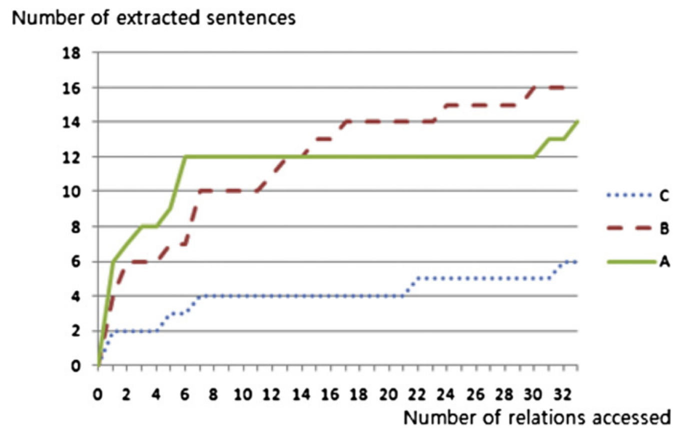

Fig. 12. Number of sentences extracted in 9% K-map.

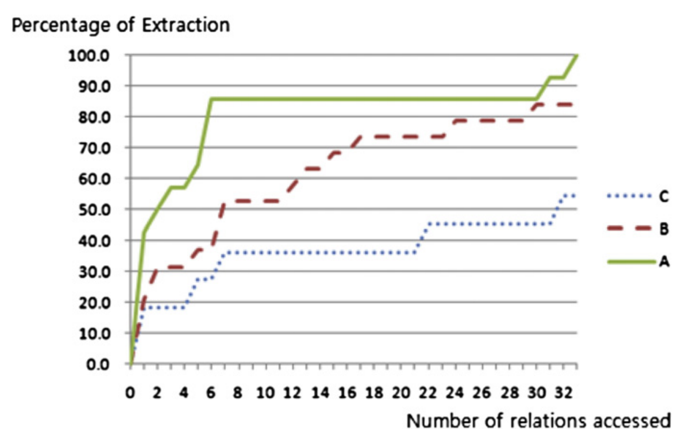

Fig. 13. Percentage of sentences extracted in 9% K-map.

group. Reasons such as concentration on reading documents for long time periods or the learning curve of the K-map model could influence the learning performance. Further experiments are needed to find out the effects of time constraints.

4.2.3. Comparison between amount of information and Sentence importance level

As in the first experiment (Section 4.1), learning performance was analyzed. Idea units were extracted from participants' answer sheets and listed. The three graders selected a classification of the importance of each idea unit. The class of the idea unit was decided according to the class chosen by the majority of graders.

Criteria for scoring sentences are similar to what appears in Section 4.1. Basic standards are the same, but they are more specified to minimize the differences among graders. The topic was the American scientist Carl Sagan. The document set consisted of three articles from three different Internet Web sites, Wikipedia (4370 words), Crystal Link (1746 words), and Imdb (603 words). All three documents described Carl Sagan's life.

The classifications were:

A: An idea unit that describes things that are directly related to the topic, so that it is thought to be highly informative.
B: An idea unit that describes things that are directly related to the topic, but is not thought to be highly informative.
C: An idea unit that describes things that are not related to the topic, so that it is thought to be not informative.

4.2.4. Comparison of the amount of not-important information

The results show that the two groups have different amounts of information not related to the main ideas of the topic. Since idea units in class C can be considered as not informative, the average number of C class idea units was compared. Table 3 shows statistical data such as mean number of C class idea units, standard deviation, and sample size.

Table 1
Results of free recall test.

Group	Observed mean	SD	Sample size
Document group	28.47	10.44	17
Map group	29.00	10.87	17
Total	28.74	10.50	34

Table 2
Comparison of the number of idea units between map and document groups.

Group	Observed mean	SD	Sample size
Document group	15.82	5.29	17
Map group	15.59	5.64	17
Total	15.71	5.39	34

Table 3
Comparison of the amount of non-related information.

Group	Observed mean	SD	Sample size
Document group	3.24	2.44	17
Map group	0.82	1.29	17
Total	2.03	2.28	34

Since normality test does not indicate that data from the Map group do not have normality ($p < 0.005$), non-parametric test (Mann-Whitney test) was applied to check whether there is a difference in the amount of unrelated information. The result indicates there is a difference ($p = 0.0031$). The results show that the Document group learned more non-important information than did the Map group.

5. Discussion

The experiments analyzed the performance of the domain knowledge map (K-map) model. From the first set of experiments (Section 4.1), we can see that K-map filters sentences that are not important to the understanding of the contents. The user can have access to at least 70% of the sentences classified as important. This percentage can be increased by having more concepts in K-map. Since parts that are not important were filtered out of the map, the user can reduce the time needed to grasp main ideas.

The results in Section 4.2.2 show that Document group and Map group can recall the same amount of information. The results of Section 4.2.4 show that there is a difference in the amount of unrelated information between the two groups. The results show that the Document group has more unrelated information than the Map group does, and therefore the Map group has more of the important information than the Document group does, given the fact that Document group and Map group have the same amount of information. The experiment results show the advantage of K-map in terms of the quality of the extracted information.

Once participants finished writing answer sheets in the experiment, they were encouraged to describe the disadvantages of K-map from their experiences. Their description of the disadvantages was collected for critical analysis of usability and for further research of the model. Five issues are discussed below.

First, the same sentence appears repeatedly in many relations. For example, if a sentence has a word phrase “stock market bubble” and K-map has the words stock, market, and bubble for its keywords, this sentence will appear in at least three relations, which are stock market, stock-bubble, and market bubble. Participants mentioned that they lost interest in learning when they saw sentences that repeatedly appeared. To reduce the inefficiency, the user needs to be informed which sentence was accessed before and which is new, so that the user does not have to read the same sentences again. One method is assigning each sentence to one relation, but here we need to have a reasonable reason why this sentence has to be in this relation. Another method is using a color code. The system stores the information regarding which sentences the user already accessed and changes the color of the sentences to inform the user that the sentence was already read. Another method can give already-accessed-sentences lower rankings when they appear in other relations and therefore position them further down in the list of sentences, so that the user can distinguish these sentences.

Second, since there are many relations on the map, some users mentioned that they clicked the same relation twice. This problem can be solved by letting them know what relations were already viewed by using different colors, similar to the second solution in the first issue.

Third, some users mentioned that although they read sentences of a relation, they sometimes barely grasped the real relationship between the two concepts. K-map only displays sentences that have chosen keywords, which are placed at both ends of a relation. It does not really describe what the relationship really is. K-map just tries to help users infer relationships by showing related sentences. This issue is considered a more difficult problem to solve. The words that represent the relationship will be mostly verbs. One option is automatic relation label selection. If extracting a relation label is successful, a proposition will be made in noun-verb-noun form, which makes K-map have concept map features.

Fourth, when a user clicks a relation, several sentences can be viewed. But there is no relationship among them. Having no predefined labeled relation sometimes made it difficult for participants to read the sentences, because there was no context among sentences. This is one of the shortcomings of learning only from K-map. The user is encouraged to first learn broad and overall knowledge about the domain and then to choose appropriate documents based on the knowledge they learned from the map and to reapply the K-map process.

Fifth, some users mentioned that some keywords appearing on the map did not look like keywords for the domain. This issue appears because the system automatically extracted keywords and thus the accuracy of these keywords is lower than that of keywords manually selected by domain experts. However, K-map can improve as keyword extraction techniques develop.

6. Conclusion

The work presented a method for automatic creation of knowledge maps. The method was analyzed using the K-map Tool for information recall and for identification of important ideas. The experiments performed show the number of relations required to identify the important ideas in the text. Furthermore, according to the experiments, K-map provides a mechanism with which to distinguish the more important sentences. Therefore, K-maps show promise as a tool for e-learning environments.

K-map offers multiple benefits, especially when used in an e-learning platform. A user can see the key concepts in a domain and can identify which concepts are strongly related to others. A user reading sentences can directly access a relevant document from a certain sentence. Future work includes analyzing whether 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. Exploring a knowledge domain using K-map can help the user see the holistic picture. Allowing the user to choose relations based on keywords can promote selective learning about the domain, which is hardly possible when learning from text.

References

Amer, A. A. (1994). The effect of knowledge-map and underlining training on the reading comprehension of scientific texts. *English for Specific Purposes, 13*(1), 35–45.
Ausubel, D. (1968). *Educational psychology: A cognitive approach*. New York: Holt, Rinehart and Winston.
Böhm, K., Heyer, G., Quasthoff, U., & Wolff, C. (2002). Topic map generation using text mining. *Journal of Universal Computer Science, 8*(6), 623–633.
Cañas, A. J., Carff, R., Hill, G., Carvalho, M., Arguedas, M., Eskridge, T. C., et al. (2005). Concept maps: integrating knowledge and information visualization. *Lecture Notes in Computer Science, 3426/2005*, 181–184.
Carvalho, M. R., Hewett, R., & Cañas, A. J. (2001). Enhancing web searches from concept map-based knowledge models. In *Proceedings of multi-conference on systems, cybernetics and informatics* (pp. 69–73).
Chen, N. S., Kinshuk, Wei, C. W., & Chen, H. J. (2008). Mining e-learning domain concept map from academic articles. *Computers & Education, 50*(3), 1009–1021.
Chen, Y., & Xia, H. (2009). Research on the auto-construction methods of concept map. In *Proceedings of international conference on intelligent human-machine systems and cybernetics*.
Chu, K. K., Lee, C. I., & Tsai, R. S. (2011). Ontology technology to assist learners' navigation in the concept map learning system. *Expert Systems with Applications, 38*, 11293–11299.
Clark, H. H., & Wasow, T. (1998). Repeating words in spontaneous speech. *Cognitive Psychology, 37*, 201–242.
Czerniak, M. C. (1998). The effect of collaborative concept mapping on elementary preservice teachers' anxiety, efficacy, and achievement in physical science. *Journal of Science Teacher Education, 9*(4), 303–320.
Dansereau, D. F., Dees, S. M., & Simpson, D. D. (1994). Cognitive modularity: implications for counseling and the representation of personal issues. *Journal of Counseling Psychology, 41*(4), 513–523.
Dicheva, D., & Dichev, C. (2006). TM4L: creating and browsing educational topic maps. *British Journal of Educational Technology, 37*(3), 391–404.
Gomez, A., Morenoa, A., Pazosa, J., & Sierra-Alonso, A. (2000). Knowledge maps: an essential technique for conceptualization. *Data & Knowledge Engineering, 33*, 169–190.
Hall, R. H., Dansereau, D. F., & Skaggs, L. P. (1992). Knowledge maps and the presentation of related information domains. *Journal of Experimental Education, 61*(1), 5–18.
Hall, R. H., & O'Donnell, A. M. (1996). Cognitive and affective outcomes of learning from knowledge maps. *Contemporary Educational Psychology, 21*(1), 94–101.
Holley, C. D., & Dansereau, D. F. (1984). *Spatial learning strategies: Techniques, applications, and related issues*. New York, NY: Academic Press.
Hou, X., Ong, S. K., Nee, A. Y. C., Zhang, X. T., & Liu, W. J. (2011). GRAONTO: a graph-based approach for automatic construction of domain ontology. *Expert Systems with Applications, 38*, 11958–11975.
ISO/IEC 13250. (2003). *Topic maps*.
Jia, H., Wang, M., Ran, W., Yang, S. J. H., Liao, J., & Chiu, D. K. W. (2011). Design of a performance-oriented workplace e-learning system using ontology. *Expert Systems with Applications, 38*, 3372–3382.
Li, J. J., & Choi, K. S. (1997). Corpus-based Chinese-Korean abstracting translation system. *Proceedings of IJCAI, 1997*, 972–976.
McClure, J. R., Sonak, B., & Suen, H. K. (1999). Concept map assessment of classroom learning: reliability, validity, and logistical practicality. *Journal of Research in Science Teaching, 36*(4), 475–492.
Nesbit, J., & Adesope, O. (2005). Dynamic concept maps. In *Proceedings of world conference on educational multimedia, hypermedia and telecommunications* (pp. 4323–4329).
Novak, J. D., & Musonda, D. (1991). A twelve-year longitudinal study of science concept learning. *American Educational Research Journal, 28*, 117–153.
O'Donnell, A. M., Dansereau, D. F., & Hall, R. H. (2002). Knowledge maps as scaffolds for cognitive processing. *Educational Psychology Review, 14*(1), 71–86.
Pepper, S. (2000). The TAO of topic maps. In *Proceedings of XML Europe*.
Porter, M. F. (1980). An algorithm for suffix stripping. *Automated Library and Information Systems, 14*(3), 130–137.
Rewey, K. L., Dansereau, D. F., Skaggs, L. P., Hall, R. H., & Pitre, U. (1989). Effects of scripted cooperation and knowledge maps on the recall of technical material. *Journal of Educational Psychology, 81*, 604–609.
Ruiz-Martínez, J. M., Valencia-García, R., Fernández-Breis, J. T., García-Sánchez, F., & Martínez-Béjar, R. (2011). Ontology learning from biomedical natural language documents using UMLS. *Expert Systems with Applications, 38*, 12365–12378.
Salton, G., & Buckley, C. (1988). Term-weighting approaches in automatic text retrieval. *Information Processing & Management*, 513–523.
Stoyanova, N., & Kommers, P. (2002). Concept mapping as a medium of shared cognition in computer-supported collaborative problem solving. *Journal of Interactive Learning Research, 13*(1), 111–134.
Tseng, S. S., Sue, P. C., Su, J. M., Weng, J. F., & Tsai, W. N. (2007). A new approach for constructing the concept map. *Computers & Education, 49*(3), 691–707.
Urfat, N., & Korhan, G. (2009). Relation extraction among learning concepts in intelligent tutoring systems. In *Proceedings of application of information and communication technologies*.
Wang, Y. H., & Liao, H. C. (2011). Data mining for adaptive learning in a TESL-based e-learning system. *Expert Systems with Applications, 38*, 6480–6485.
Wang, S. L., & Wu, C. Y. (2011). Application of context-aware and personalized recommendation to implement an adaptive ubiquitous learning system. *Expert Systems with Applications, 38*, 10831–10838.
Wiegmann, D. A., Dansereau, D. F., McCagg, E. C., Rewey, K. L., & Pitre, U. (1992). Effects of knowledge map characteristics on information processing. *Contemporary Educational Psychology, 17*(2), 136–155.