Concept Maps: Theory, Methodology, Technology
Proc. of the Fifth Int. Conference on Concept Mapping
Valletta, Malta 2012

USING AUTOMATICALLY GENERATED CONCEPT MAPS FOR DOCUMENT UNDERSTANDING: A HUMAN SUBJECTS EXPERIMENT

Alejandro Valerio & David Leake, Indiana University, U.S.A
Alberto J. Cañas, Institute for Human and Machine Cognition (IHMC), U.S.A
Email: {avalerio, leake}@cs.indiana.edu

Abstract. Concept maps present information in a concise and easily-understood form. Consequently, concept maps of documents are a useful vehicle for summarizing the documents' contents. Concept-map based summaries can in turn be used as the basis for browsable indices to help guide navigation through documents to find material of interest. However, using concept maps in this role depends on the ability to generate a concept map for each document. Especially for large document sets, constructing the needed concept maps by hand may be prohibitively expensive. This paper addresses this problem with an algorithm to automatically generate concept map fragments to aid in document understanding. The algorithm was evaluated with a human subjects study assessing the value of its results to facilitate locating and understanding portions of a document of interest. The study compared subjects' speed and accuracy in answering questions about material contained in a document when using only the document, a manually constructed concept map, or an automatically-generated concept map fragment. The study showed that providing automatically generated concept maps improved user speed while retaining accuracy, for documents whose size enabled capturing most key concepts in a single concept map..

1 Introduction

Concept maps (Novak & Gowin, 1984) provide a useful medium for organizing document information in a concise and accessible form. When captured with electronic concept mapping tools supporting the annotation of concepts with documents, concept maps can serve as a tool to help humans visualize document sets and access documents in an efficient way. Concept maps may also be used to summarize the content of single documents, to facilitate understanding of key concepts and relationships within a document. Those document concept maps may in turn be used as an index for related text passages within the document, to facilitate access to material of interest within the document. Concept map based document summaries and indices could be applied in many contexts to aid document assessment and location of useful material. For example, in training contexts, a concept-map-based summary of a manual could help a user to decide its relevance to a new problem. Likewise, a concept-map-based index captured by an electronic concept mapping tool could help the user find important information within the manual. In the context of intelligence analysis, analysts must identify relevant documents within a huge pool of candidates, without time to read each one. In this case, a concept-map-based summary could assist in rapidly identifying relevant documents, and, when documents are found to be relevant, at identifying the areas of the document to scrutinize for useful information. Here the concept map summary of the main concepts within a document could enable rapidly forming a "big picture" understanding of the document, for determining whether a document merits further examination. If the concept map summary is annotated with information about where within the document may be found the material associated to each of the concept map's concepts, it can also provide an index for finding passages about relevant concepts.

Unfortunately, the feasibility of using document-based concept-maps to support document understanding has been impeded by a serious knowledge modeling bottleneck. A manual approach to generating the maps, in which a human analyzes each document to generate the associated concept map is prohibitively expensive, especially if the goal—as in intelligence analysis—is to alleviate the need for analysts to read each document. Automated concept-map generation tools for this task could significantly help overcome this obstacle. Previously we investigated automated methods to help bridge the gap between documents and concept maps by generating concept map fragments (Valerio and Leake, 2006), and in this paper we examine how an extension of such methods can benefit human document understanding.

The difficulty of automated document understanding is well known, and full automated document understanding remains beyond the state of the art in AI. Consequently, it is important to note that our methods for generating concept maps from documents are not intended to achieve full document understanding. Instead, our aim is to use natural language processing and text mining techniques in order to reach an intermediate level of analysis, deeper and more refined than keyword-based methods, and to present the results in the form of concept maps. We believe that the automatic generation of concept maps of sufficient quality to support document summarization and indexing for training and analysis is feasible. Such maps can be useful even without capturing all the information contained in a document, provided that they capture a sufficient set of the top-level concepts and relations to (1) aid in a quick skim of the document to see if it may be relevant to a topic of interest to the user, and, if the document is potentially relevant, to (2) to serve as a conceptual summary/index to assist in a basic understanding of the document and of where in the document the user can find useful information.. Thus our criterion for success is whether the methods are sufficient for aiding a human understander.

This paper presents research on the automated construction of concept map fragments and evaluates its value for helping humans to (1) determine the relevance of textual documents, and (2) understand key points faster and more accurately than relying on the document alone. For this task, the central natural language processing requirement is accurate labeling of concepts and linking phrases. The paper presents an algorithm we have developed for this task and its evaluation with a human-subjects study. The study shows that providing subjects with the automatically-generated concept map can improve subjects' speed at answering questions about the content of a document. Consequently, we believe that our algorithm a promising step towards automated concept map generation for this type of support role. In addition, we believe that our algorithm may prove a useful step towards more sophisticated automatic concept map generation from documents.

2 Concept maps as a representation of text documents

Concept maps present information in a concise and easily-understood form. Consequently, a concept map of the content of a document is a useful vehicle for summarizing what a document is about. In addition, if the concept map is annotated with pointers to relevant passages in the text, the concept map can provide a browsable index to help guide navigation through documents to find material of interest. However, despite the relative ease of concept map construction, it may be time consuming to determine the right content for a map or its related resources. This problem is further aggravated when considering the large amounts of information available in document collections.

In order to assist users in the creation of better concept maps, as well as finding relevant resources with which to enrich them, researchers have previously aimed to develop tools for automatic support of concept map generation from source documents, as summarized in (Kowata, Cury, & Silva Boeres, 2010). Most of these approaches use natural language processing techniques at the core of their algorithms. Such techniques are commonly used to automatically extract information from unstructured text documents, for example during automatic population of ontologies (Alani et al., 2003). Syntactic parsers are at the core of these automatic processes and provide deep (Charniak & Johnson, 2005) or shallow structural analysis (Abney, 1996) of natural language sentences.

The work in (Alves, Pereira, & Cardoso 2001) uses WordNet to extract a hierarchy of nouns from the text (building an initial list of concepts), followed by several iterations with a human user to obtain relationships between pairs of concepts and find the linking phrases. Another approach relies on a predefined list of domain specific concepts provided by an expert (Clariana & Koul, 2004). This method considers two concepts to be related if they occur in the same sentence, but does not suggest possible linking phrases. These two approaches share a common drawback of requiring user input during their processing.

Another alternative focuses on word sense disambiguation (Rajaraman, & Tan, 2002), using the meaning of nouns and verbs to search for Noun-Verb-Noun structures in the sentences, which become the *concept - linking phrase - concept* relations. In (Leskovec, Grobelnik, & Milic-Frayling, 2004) a semantic graph is constructed from a document which is used for automatic document summarization. In (Zouaq & Nkambou, 2008), domain ontologies are built from text for educational purposes using a semi-automatic framework that produces a domain concept map from text, which is then used to derive a domain ontology.

The text mining and natural language processing techniques outlined above can be used to effectively harness information from documents. Through them, our approach exploits the syntax of document sentences to identify the topic, objects and relations described in such documents. This information is then used to generate a structural representation of the document content.

Our approach uses the syntactic structure of the sentences and their dependency information to find relations between the words. The relations are not retrieved from predefined ontologies, but are generated from the document itself. This enables our approach to be applied in any domain and makes the results potentially more sensitive to the intentions of the document author (which is important to the document summarization task). For example, even if two concepts are related in a particular ontology, the author of a document might have intentionally ignored that relation, because it did not correspond to the desired level of abstraction; in such circumstances, our approach would not include the relationship in the document description. In addition, our algorithm produces concepts based on a sentence parser trained to recognize specific syntactic structures rather than producing concepts based on individual words, making the concept labels more complete.

3 Automatic Generation of Concept Maps to Assist in Document Understanding

We have presented our algorithm for automatic generation of concept maps in previous work (Valerio, Leake, & Cañas, 2008a). In this section, we present an overview of the algorithm (as seen in Figure 1) and describe the modifications necessary to generate concept maps for the specific task of assisting in document understanding.

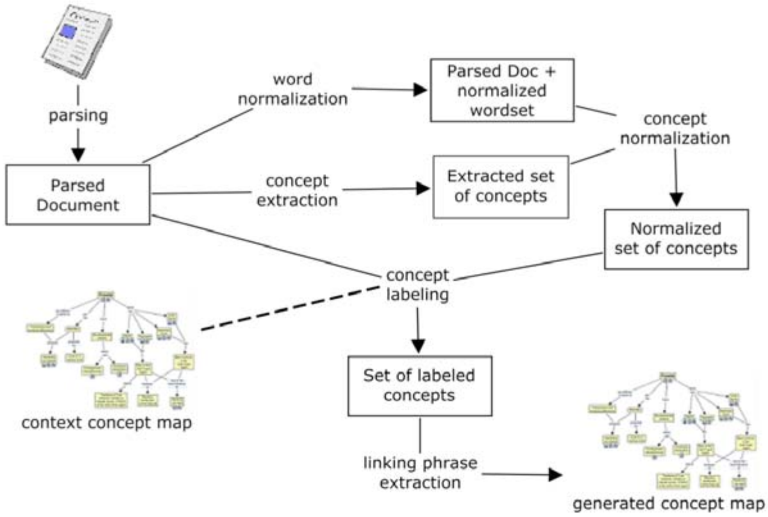


Figure 1. Procedure to construct a concept map fragment automatically from a document.

We assume the input document is "well-written" and that it contains the description of a concept or set of concepts. Each step in the algorithm is responsible for generating different parts of the final concept map, such as a list of concepts or linking phrases. Where appropriate, we use available natural language processing tools in order to generate the parse trees from the incoming document that are needed to construct a concept map. Our algorithm refines the quality of the output at each step to ensure the final output of the algorithm is not merely a graph of interconnected nodes. Instead, the goal is to generate a hierarchical structure with a root node that matches the topic of the document and that shares the characteristics of well-constructed concept maps.

In its simplest form, a concept map is a set of propositions where each proposition is a triple of the form concept - linking phrase - concept. We assume that concepts can be found in the noun phrases of the text and that linking phrases correspond to verb and prepositional phrases. After parsing the document, the resulting parse trees are used as input for the next steps in the algorithm. First, the algorithm takes all the nouns from the parse trees and performs word normalization, by grouping words considered to be equivalent based on the lexical root of the word, or words that are found to be synonyms.

The concept extraction step takes the noun phrases from the parse trees and tags these as potential concepts. Concept normalization combines the results of word normalization and concept extraction by removing morphological variations on words and performing anaphora resolution. The result is a list of abstract concepts, each composed of a set of noun phrases that are considered to be equivalent. The algorithm next assigns a concept label to each of the abstract concepts, with the label selected from the concept's associated set of noun phrases. By default, it selects the longest noun phrase, assuming it will be the most descriptive label. However, in this step it is possible to consider the input of context in the form of other concept maps, which may help choose a more appropriate label, based on the similarity of the topic of the context concept map and the concept labels contained within.

Finally, the algorithm finds the relationships between the concepts by finding the verb and prepositional phrases in the parse trees that connect the concepts that it has identified and tagged in previous steps. After finding the linking phrases that connect the concepts, it is left with a set of triples that can be used to construct a concept map that summarizes the information contained in the input document.

For the purpose of our human document understanding experiment, we tuned our algorithm to generate maps with certain characteristics resembling those of concept maps built by humans. Heuristics based on previous work (Valerio, Leake, & Cañas, 2008b) made the automatically generated maps resemble human maps in terms of the average number of concepts and linking phrases, cross-links, and overall hierarchy and taxonomy of the concept map.

The characteristics of concept maps generated by our algorithm depend on the length of the source documents, which has a direct effect on the number of concepts identified during the concept extraction phase. Because we want our generated concept maps to resemble human-made concept maps, and for our concept maps to be readable for the subjects, we limit the number of concepts in the automatically generated concept maps to 30, as well as performing some trimming of the final concept map, removing leaf nodes with no siblings and isolated sections with only 1 or 2 concepts. As a result, it is possible that some information is lost in the automatically generated concept maps of the longer documents, which may be information needed in order to correctly answer the reading comprehension question.

Note that these limitations are not needed when the automatically generated concept maps will be processed by some other automated procedure, when human readability is not a primary concern. However, in the context of a reading comprehension experiment, more information in the concept map might generate too many visual elements for the subjects to process, reducing their reading comprehension capabilities. Furthermore, our main concern in this experiment is not to provide complete document summaries, but rather general overviews that can be quickly understood and that can give a subject a general idea of the topic of the document and the relation between the most important concepts present in the document.

4 Overview of the Approach

We hypothesize that providing concept map fragments generated automatically by our algorithm to users will enable them determine the relevance of these documents and to help them understand key points faster and more accurately than by relying on the document alone. To test this hypothesis we designed a human subjects reading comprehension experiment, in which users are presented with a reading comprehension task. In this task, users must answer questions about a document, given one of three different representations of the same document: the document in its original form, a human-made concept map based on the document, and an automatically generated concept map from the document. The experiment then measures their reading comprehension skills (in terms of both accuracy and speed) with each of the three representations. The test evaluates the accuracy and efficiency of subjects when answering questions about the content of the document. This evaluation not only tests the ability of the concept maps to help humans understand documents, but also gives evidence of the quality of the concept map that is produced.

Previous experiments have tested the usefulness of concept maps as a means to assist users in the reading comprehension of texts of English as a second language (Dias 2010, Rosenberg & Saif, 2010). In these studies the researchers found a clear improvement in students' reading comprehension skills, which they attribute to the fact that concept maps can help a student to visualize how ideas are organized by the author of a text. Concept maps were also evaluated as a comprehension aid for students with hearing impairments (Castillo, Mosquera, & Palacios, 2010). In this study, the subjects were able to raise their reading comprehension levels while increasing their attention spans on a particular subject.

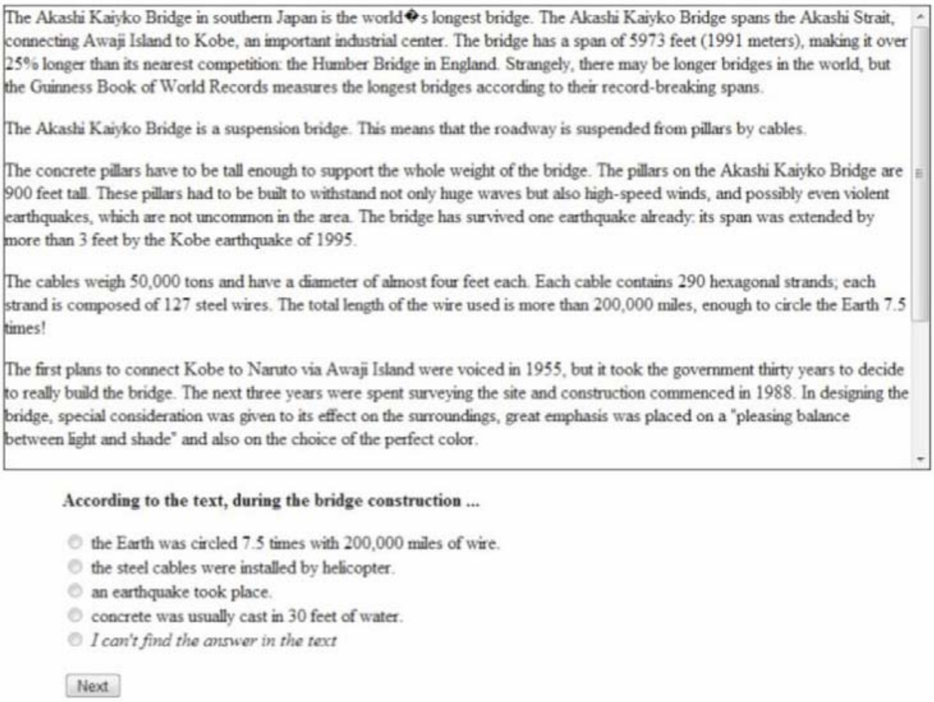
Another experiment presented a group of students with three different types of concept mapping tasks in order to measure in which one reading comprehension levels increased the most (Chang, Sung, & Chen, 2010). In the first task students were given concept maps constructed by teachers. These concept maps purposefully contained erroneous information, which the students were required to correct. In the second task, students were presented with a complete and correct concept map on a particular subject. Over the course of the term, the provided concept maps contained less information, forcing the students to complete the concept map. By the end of the term, the students were expected to construct their own concept maps. The results of this experiment showed that while students in all three conditions showed an overall improvement in their reading comprehension skills, those that were required to correct the concept maps achieved the best overall results. The proposed explanation was that map correction requires the students to apply critical and analytical thinking of a subject which leads to deeper processing of the new information.

Our experiment studies the use of concept maps to assist users with general reading comprehension tasks. Although there exist several methods for representing and summarizing text in the context of reading comprehension tasks (Kozminsky, Nathan, & Kozminsky 2010), it is not our intention to compare concept maps with other types of representations. Rather, we are interested in determining if the concept maps generated with our algorithm are good enough to help users to better understand a document, and if they are comparable to human generated concept maps.

This experiment is different from the previous tasks described in that our concept maps are automatically constructed, while all of the above experiments use human-made concept maps or other forms of structured representations of documents. We are aware that our automatic concept map construction may not accurately represent some relationships (e.g., those that are given by the order in which a particular text is written), and are testing whether our algorithm is able to extract enough information for the resulting concept map to be useful in helping users understand the text and correctly answer the reading comprehension questions.

5 Experimental setup and results

Subjects in our experiment complete a series of reading comprehension questions. For each question, subjects are presented with the necessary information in order to answer the question. This information may be presented in one of three different representations: a plain text document, a human made concept map, or a concept map that was automatically generated with our algorithm. Both concept maps are constructed using their corresponding plain text document as source. Subjects do no know whether the concept maps they see are human made or automatically generated. Figure 2 illustrates a question with information provided as a plain text document, and Figure 3 shows a question with information provided by a concept map.


The Akashi Kaiyko Bridge in southern Japan is the world�s longest bridge. The Akashi Kaiyko Bridge spans the Akashi Strait, connecting Awaji Island to Kobe, an important industrial center. The bridge has a span of 5973 feet (1991 meters), making it over 25% longer than its nearest competition: the Humber Bridge in England. Strangely, there may be longer bridges in the world, but the Guinness Book of World Records measures the longest bridges according to their record-breaking spans.

The Akashi Kaiyko Bridge is a suspension bridge. This means that the roadway is suspended from pillars by cables.

The concrete pillars have to be tall enough to support the whole weight of the bridge. The pillars on the Akashi Kaiyko Bridge are 900 feet tall. These pillars had to be built to withstand not only huge waves but also high-speed winds, and possibly even violent earthquakes, which are not uncommon in the area. The bridge has survived one earthquake already: its span was extended by more than 3 feet by the Kobe earthquake of 1995.

The cables weigh 50,000 tons and have a diameter of almost four feet each. Each cable contains 290 hexagonal strands; each strand is composed of 127 steel wires. The total length of the wire used is more than 200,000 miles, enough to circle the Earth 7.5 times!

The first plans to connect Kobe to Naruto via Awaji Island were voiced in 1955, but it took the government thirty years to decide to really build the bridge. The next three years were spent surveying the site and construction commenced in 1988. In designing the bridge, special consideration was given to its effect on the surroundings, great emphasis was placed on a "pleasing balance between light and shade" and also on the choice of the perfect color.

According to the text, during the bridge construction ...

- the Earth was circled 7.5 times with 200,000 miles of wire.
- the steel cables were installed by helicopter.
- an earthquake took place.
- concrete was usually cast in 30 feet of water.
- *I can't find the answer in the text*

Next

Figure 2. Human subject reading comprehension experiment: plain text document question.

The text and corresponding reading comprehension questions are taken from standard Q&A tests designed for middle school students at the 7th and 8th grade level. The subjects of the experiment were Indiana University students, both undergraduate and graduate, who were compensated for their time. In total, there were 16 participating subjects. Each subject answered a total of 60 reading comprehension questions, with 20 questions for each different representation: plain text, human made concept map, and automatically generated concept map. The subjects were divided into three distinct sets such that each question was given in a different representation to each set of users, as shown in Table 1.

For each question, we recorded both the accuracy of the response (as given by the answer key in the reading comprehension test) and the time it took each subject to answer the question. We added a *"Could not find answer"* response as an option in all of the reading comprehension questions. This is of particular importance in the case of the concept maps: As we have already mentioned, it is possible for some information to be left out during concept map generation. A choice of *"Could not find answer"* by the subjects was treated as an incorrect answer to the question. Our experimental objective is to measure whether concept maps can assist in helping user's reading comprehension skills. If our approach fails because the concept map did not contain the relevant information, or because the subject was not able to understand the concept map, the algorithm has failed its objective.

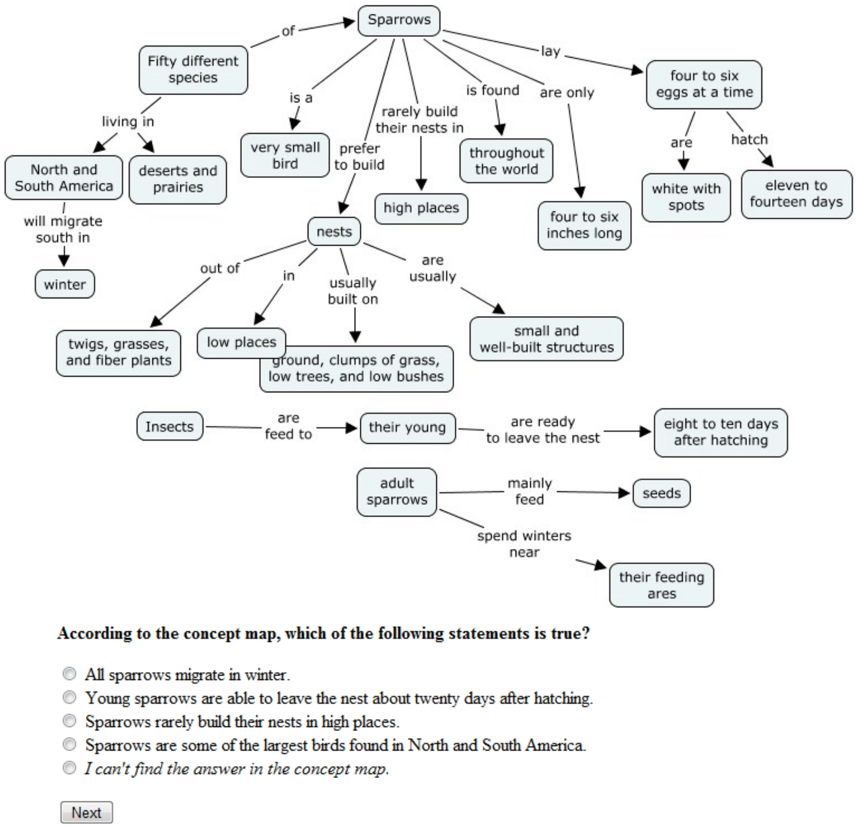


Figure 3. Human subject reading comprehension experiment: concept map question.

Our analysis compares both the accuracy of the subjects' answers and the amount of time they needed in order to answer a question in each of the three representations of the information. Additionally, we clustered the documents according to their length: 1 to 2 paragraphs, or 4 to 5 paragraphs, in order to measure the effect of concept map trimming on reading comprehension. The main results of our experiment are shown in Table 2 and Figure 4.

	Text	**Human-made**	**Automatically generated**
Set 1	Questions 1 - 20	Questions 21 - 40	Questions 41 – 60
Set 2	Questions 21 - 40	Questions 41 - 60	Questions 1 – 20
Set 3	Questions 41 - 60	Questions 1 - 20	Questions 21 – 40

Table 1: Human subject experiment: question distribution for sets of subjects.

Overall, we see that the use of concept maps does not increase the accuracy of the subjects' reading comprehension skills, and in fact, provides worse accuracy than plain text. However, we note the trade-off between accuracy and speed. While the use of automatically generated concept maps resulted in an overall average decrease in accuracy of 16% when compared to the text document versions, it also accounted for an average decrease of 43% of the time that subjects took to answer questions.

	Accuracy (1-2 parags)	Time (1-2 parags)	Accuracy (4-5 parags)	Time (4-5 parags)
Text	86%	41 seconds	68%	61 seconds
Human-made	81%	23 seconds	53%	38 seconds
Automatically generated	75%	24 seconds	47%	32 seconds

Table 2: Human subject experiment results

A closer look suggests the importance of document length. The difference in the accuracy results shown between the 1-2 and 4-5 paragraph documents suggests that there is indeed a loss of information in concept maps generated automatically from larger documents, due to the limit imposed on the number of concepts in the automatically generated concept map. While in the 1-2 paragraph documents accuracy degrades by an average of 11% compared to text, the degradation is on average 21% for the 4-5 paragraph documents. The average time decrease of 43% is shared by both types of documents, suggesting that concept maps are a more effective way in which to quickly convey information to users.

When considering the difference between the two types of concept maps, our goal was to automatically generate concept maps that were similar to human-made concept maps. It appears that we succeeded in this goal for the 1-2 paragraph documents, for which the performance of both types of concept maps showed no significant difference. However, we found a trade-off between accuracy and speed for the 4-5 paragraph documents, where the human-made concept maps allowed subjects to answer the questions with an average of 11% more accuracy, while the automatically generated concept maps allowed them to answer the questions an average of 16% faster than the human-made concept maps.

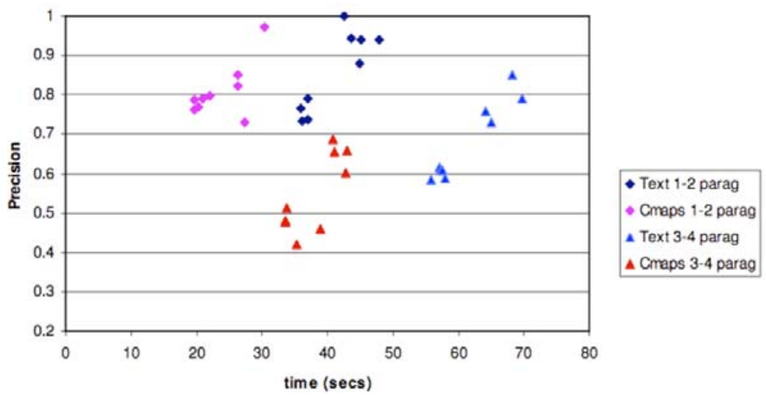


Figure 4. Scatter plot of average time and accuracy per user on the reading comprehension task.

From the results we conclude that providing concept maps did not assist users in improving accuracy compared to providing the initial text. This is more evident for longer documents, in which restrictions on the number of concepts decreased the information provided. However, the accuracy results still indicate that the major concepts of the document were correctly recognized and were available to the subjects, to allow for an understanding of the overall topic of the document. In addition, when the needed information is present in the concept map, subjects are able to find the correct information much faster than in the source document. Therefore, we conclude that our automatically generated concept maps help users improve their performance in document understanding tasks, in terms of the speed in which they can find the required information.

6 Summary

Our experiments investigated whether automatically generated concept maps, produced by our algorithm, could assist humans in a document understanding task. We showed that our concept maps allow users to substantially improve their reading comprehension skills in terms of speed. In particular, we found that for smaller documents with fewer concepts, almost every concept is reflected in the constructed concept map, giving users comparable accuracy in reading comprehension in addition to a significant speedup. For larger documents, the accuracy of the answers decreased but the speedup was maintained. We conclude that our algorithm is promising for tasks where the size of documents is small or when the response time is of major importance. The key to the usefulness of this method is to highlight information which might otherwise be hidden in a large document library. In general, we have found that the automatically produced concept maps are of similar quality to those authored by humans, and we believe the steps taken in this research are an important step in the direction of enhancing automatic processing of documents into concept maps.

7 References

Abney, S. P. (1996). Part-of-Speech Tagging and Partial Parsing. In Church, K., Young, S., and Bloothooft, G., editors, Corpus-Based Methods in Language and Speech. Kluwer Academic Publishers.

Alani, H., Kim, S., Millard, D. E., Weal, M. J., Hall, W., Lewis, P. H., & Shadbolt, N. R. (2003). Automatic Ontology-Based Knowledge Extraction from Web Documents. IEEE Intelligent Systems, 18(1):14–21.

Alves, A. O., Pereira, F. C., and Cardoso, A. (2001). Automatic Reading and Learning from Text. In Proceedings of the International Symposium on Artificial Intelligence (ISAI'2001), pages 302–310.

Chang, K.-E., Sung, Y.-T., and Chen, I.-D. (2002). The Effect of Concept Mapping to Enhance Text Comprehension and Summarization. Journal of Experimental Education.

Charniak, E. and Johnson, M. (2005). Coarse-to-fine n-best Parsing and MaxEnt Discriminative Reraking. In Proceedings of the 43rd Annual Meeting of the ACL.

Clariana, R. B. and Koul, R. (2004). A Computer-Based Approach for Translating Text into Concept Map-like Representations. In A. J. Cañas, J. D. Novak & F. M. González (Eds.), *Concept Maps: Theory, Methodology, Technology. Proc. of the 1st Int. Conf. on Concept Mapping*, Pamplona, Spain.

Dias, R. (2010). Concept Map: A Strategy for Enhancing Reading Comprehension in English as L2. In J. Sánchez, A. J. Cañas, J. D. Novak (Eds.), *Concept Maps: Making Learning Meaningful, Proc. of the Fourth Int. Conference on Concept Mapping,* Viña del Mar, Chile: Universidad de Chile.

Kowata J. H., Cury D., & Silva Boeres M. C. (2010). A Review of Semi-Automatic Approaches to Build Concept Maps. In J. Sánchez, A. J. Cañas, J. D. Novak (Eds.), *Concept Maps: Making Learning Meaningful, Proc. of the Fourth Int. Conf. on Concept Mapping,* Viña del Mar, Chile: Univ. de Chile.

Kozminsky, E., Nathan, N., and Kozminsky, L. (2010). Mapping Some Design Considerations for Learning From Texts In J. Sánchez, A. J. Cañas, J. D. Novak (Eds.), *Concept Maps: Making Learning Meaningful, Proc. of the Fourth Int. Conference on Concept Mapping,* Viña del Mar, Chile: Universidad de Chile.

Leskovec, J., Grobelnik, M., and Milic-Frayling, N. (2004). Learning Sub- structures of Document Semantic Graphs for Document Summarization. In Proceedings of the 10th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining.

Novak, J. D., & Gowin, D. B. (1984). Learning How to Learn. New York: Cambridge University Press.

Peña Castillo, E. P., Mosquera, D. B., and Palacios, D. L. (2010). Concept Maps: A Tool to Improve Reading Comprehension Skills of Children with Hearing Impairments. In J. Sánchez, A. J. Cañas, J. D. Novak (Eds.), *Concept Maps: Making Learning Meaningful, Proc. of the Fourth Int. Conference on Concept Mapping,* Viña del Mar, Chile: Universidad de Chile.

Rajaraman, K. and Tan, A.-H. (2002). Knowledge Discovery from Texts: A Concept Frame Graph Approach. In Proceedings of the 11th Int. Conference on Information and Knowledge Management, pages 669–671.

Rosenberg, R. and Saif, S. (2010). Computer-Based Concept Mapping for Second Language Reading Comprehension: A Design-Based Research Study. In J. Sánchez, A. J. Cañas, J. D. Novak (Eds.), *Concept Maps: Making Learning Meaningful, Proc. of the Fourth Int. Conference on Concept Mapping,* Viña del Mar, Chile: Universidad de Chile.

Valerio, A. and Leake, D. Jump-Starting Concept Map Construction with Knowledge Extracted from Documents. Concept Maps: Theory, Methodology, Technology. In: A. J. Cañas & J. D. Novak (Eds.), *Concept Maps: Theory, Methodology, Technology. Proceedings of the Second International Conference on Concept Mapping*. San José, Costa Rica: Universidad de Costa Rica, pp. 181-188.

Valerio, A., Leake, D., and Cañas, A. J. (2008a). Associating Documents to Concept Maps in Context. In In A. J. Cañas, P. Reiska, M. Åhlberg & J. D. Novak (Eds.), *Concept Mapping: Connecting Educators, Proc. of the Third Int. Conf on Concept Mapping,* Tallinn, Estonia & Helsinki, Finland: Univ. of Tallinn.

Valerio, A., Leake, D., and Cañas, A. J. (2008b). Automatic Classification of Concept Maps Based on a Topological Ontology. In A. J. Cañas, P. Reiska, M. Åhlberg & J. D. Novak (Eds.), *Concept Mapping: Connecting Educators, Proceedings of the Third International Conference on Concept Mapping*, Tallinn, Estonia & Helsinki, Finland: University of Tallinn.

Zouaq, A. and Nkambou, R. (2008). Building Domain Ontologies from Text for Educational Purposes. In IEEE Transactions on Learning Technologies.