

Concept maps as visual interfaces to digital libraries: summarization, collaboration, and automatic generation

Rao Shen, Ryan Richardson
Computer Science Department
Virginia Tech
Blacksburg, VA 24061
{rshen, ryanr}@vt.edu

Edward A. Fox
Computer Science Department
Virginia Tech
Blacksburg, VA 24061
fox@vt.edu

ABSTRACT

There is a need to make digital libraries (DLs) become knowledge repositories instead of information providers. Concept maps as knowledge visualization tools have great potential for DL application. Through our experience with GetSmart, a DL service with support for concept maps, we found that concept maps can enhance collaborative learning and serve as overviews that summarize the main ideas of a domain. Our current work on the automatic generation of concept maps through text mining techniques is presented in this paper.

Keywords

Knowledge Visualization, Digital Library, Concept Maps

1. DIGITAL LIBRARY AND DATA-INFORMATION-KNOWLEDGE CONTINUUM

It is generally agreed by IT practitioners that there exists a continuum of data, information, and knowledge within any enterprise. Data are mostly structured, factual, and often numeric. Information is factual, but unstructured, and in many cases textual. Knowledge is inferential, abstract, and is needed to support business decisions. In addition to the IT view of the data-information-knowledge continuum, other researchers have taken a more academic view. For example, information scientists consider taxonomies, subject headings, and classification schemes as representations of knowledge. Artificial intelligence researchers have long been seeking such ways to represent human knowledge as semantic nets, logic, production systems, and frames [2]. Knowledge derives from information just as information derives from data.

For digital library researchers, there is a clear need to transform information access to knowledge creation. Instead of serving as information providers, digital libraries could become knowledge repositories by effectively categorizing, analyzing, and organizing their contents [2].

LEAVE BLANK THE LAST 2.5 cm (1") OF THE LEFT COLUMN ON THE FIRST PAGE FOR THE COPYRIGHT NOTICE.

2. KNOWLEDGE VISUALIZATION TOOL: CONCEPT MAP

The most common objective for a knowledge visualization tool is to make visible an intellectual landscape [9]. One such knowledge tool is a concept map, which is a tool for organizing knowledge and representing concepts and relationships as node-link diagrams. Concept mapping is considered a constructivist learning activity [6].

Concept maps have great potential for digital library applications. GetSmart is a digital library service with support for concept maps [7]. Though in January 2003, we began running GetSmart servers at Virginia Tech for two sections of our spring class "Multimedia, Hypertext, and Information Access" (CS4624), in the fall of 2002 all use of GetSmart was hosted by the AI Laboratory of the University of Arizona. In CS4624 this spring and a graduate class "Information Storage and Retrieval" (CS5604) last fall, each student created a number of concept maps. Students can use their maps to check their own understanding of the concepts and plan their study time, as well as aid faculty in diagnosing and correcting student misconceptions. Through our experience with GetSmart, we found that concept maps can enhance collaborative learning and serve as overviews that summarize the main ideas of a domain.

2.1 Summarization

Most of the concept maps the students created were related to the textbooks. Each student completed a concept map using GetSmart from home, for each chapter assigned. A concept map drawn by a student is shown in Figure 1.

Figure 2. Concept map drawn by a student

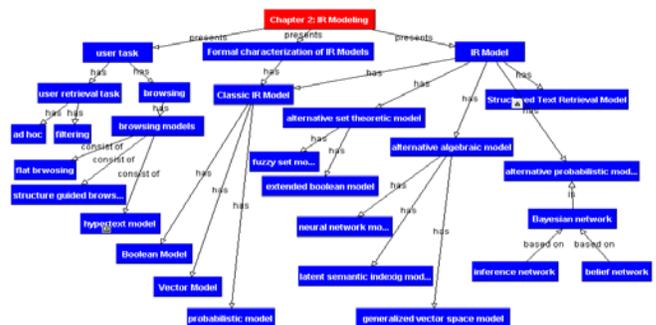


Figure 1. Concept map drawn by a student

Since the spatial properties of concept maps convey organizational information (macrostructure) that facilitates the construction of an overview (macrostructure) of a domain, the concept maps for the chapters are a summarization of the textbook as well as knowledge created by students.

During the construction of their concept maps, students used DL services and collections provided by GetSmart. To further understand concepts they identified in the chapters, students used GetSmart's metasearch interface to get related resources in DLs such as CITIDEL [1] and ACM Digital Library, and then organized those resources through concept maps. Concept maps creation as summarization, and a special form of meaningful learning helps to add knowledge layers to the DLs' collections.

2.2 Collaboration

The use of concept maps in a classroom environment has shown that concept maps are of benefit in both individual and collaborative settings [7].

Concept maps take advantage of DLs to further address the needs of learners in DLs. In CS4624 and CS5604, students' personal maps could be authorized to their group members as shared group maps, which could be completed in the distributed environments geographically and synchronized. Group maps presented in class built a communication bridge between instructor and learners and among learners. Comments and suggestions from classmates and instructor helped the group members identify misconceptions and reinforce understanding.

2.3 Automatic generation

Students created concepts maps from scratch or from templates, which could be developed by domain experts or automatically created. It is possible to have concept maps that in whole or major part automatically generated through some computational process. Students created concept maps from scratch or from templates, which could be developed by domain experts or automatically created. It is also possible to have concept maps that were, in whole or in part, automatically generated through some computational process. This was done pioneered by Gaines and Shaw [4], using term co-occurrence. More recent work was done by Saito, et al., for Japanese documents [10] and Rajaraman and Tan [8] for English documents. There are still serious problems to overcome however. A few of these are that words do not necessarily equate to concepts, that writers use synonyms which should map to the same concept, and that determining the relations between concepts is difficult to do automatically. Noun phrasers, such as the one developed at the University of Arizona, help us deal with phrases that should be considered one concept. Efforts such as Wordnet [3] help us deal with the synonymy problem. The final problem appears to be the hardest to solve. Saito et al used part-of-speech information to determine which words should be

nodes and which should be links. We are continuing this work to produce concept maps for both English and Spanish documents.

We have conducted a preliminary study, with 28 subjects, which showed that, at least for some types of questions, having access to a concept map allowed users to rank the relevance of a paper better than having the abstract alone. With this result in mind, our goal is to be able to automatically generate concept maps of sufficient quality that they can be used as a summarization and knowledge discovery tool. These could also be used to supplement other summarization techniques such as automatic text summarization and self-organizing maps.

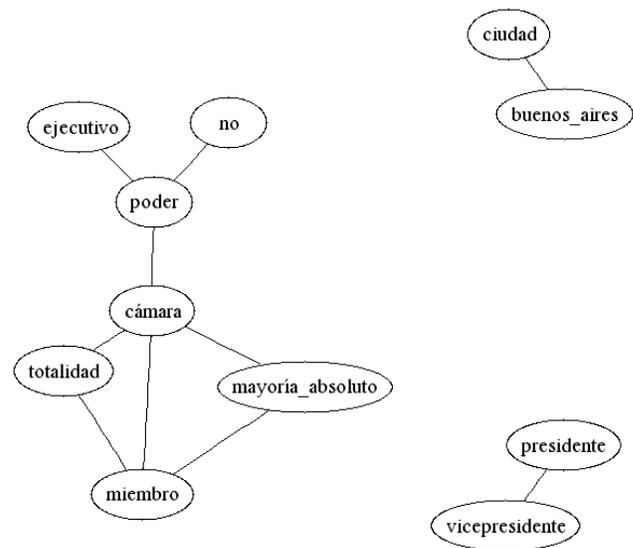


Figure 2: Automatically generated concept map of Argentine constitution

3. DISCUSSION

Concept maps as visualization tools have great potential for DL applications. Some of their supports for DLs are listed as followings.

3.1 Constructivist learning and knowledge creation

In a DL context, concept maps created by learners or automatically generated serve as summarization of a domain, which help to build a knowledge layer over collections of a DL. In a DL learning environment, concept maps take advantage of DLs to further engage both learners and instructors in an active learning approach and collaboration setting.

3.2 Browsing and searching assistant

Concept maps as a visual interface to DLs can assist users in browsing and searching.

Most visualization tools do not support more than one classification scheme for users to browse resources, and each browse scheme is based on a tree structure. Multiple taxonomies can be made available to users of DLs, so they can utilize those they find most convenient. A summary of high-dimensional collections of a DL or a retrieval result

set can be displayed as a concept map, which may provide access to multiple taxonomic perspectives and the relationships among them.

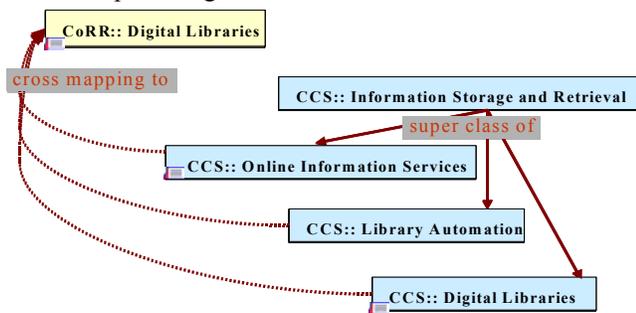


Figure 3. Searching result visualized in Concept map as cross-linked trees

Search results can be visualized in a concept map as cross-linked trees as shown in Figure 3. Each node in a concept map represents a category of a classification scheme, which can be expanded to view the details of all the documents in that particular classification scheme.

3.3 Concept modeling of DLs

The current demand from non-experts who wish to build DLs is strong and wide, and there is a need to simplify the modeling process and the rapid generation of DLs. We used concept maps to semantically and visually model DLs based on a metamodel that describe DLs using the 5S theory[5] [12].

4. FUTURE WORK

Based on our previous work with the GetSmart project, we plan to create a publicly accessible repository of concept maps to enable sharing and reuse of the knowledge represented in concept maps. A search engine component developed in CITIDEL will provide services to facilitate discovery and reuse involving the knowledge repository. We also hope to make GetSmart more portable by packaging it as an ODL component [11].

ACKNOWLEDGMENTS

This research work was founded in part by the NSF through grant DUE-0121741. We also thank the GetSmart team of the AI Lab at the University of Arizona, where GetSmart was developed.

REFERENCES

1. CITIDEL, CITIDEL Homepage, in Computing and Information Technology Interactive Digital Educational Library. 2001, Virginia Tech: Blacksburg, VA, USA.

2. Chen, H.C., Knowledge Management System. 2001, Tucson: The University of Arizona.
3. Fellbaum, C., WordNet: An electronic lexical database. 1998: The MIT Press, Cambridge, Massachusetts.
4. Gaines, B.R. and M.L.G. Shaw. Using knowledge acquisition and representation tools to support scientific communities. in AAAI'94: Proceedings of the Twelfth National Conference on Artificial Intelligence. 1994. Menlo Park, California, USA: AAAI Press/MIT Press.
5. Gonçalves, M. A., Fox, E. A.: 5SL – a language for declarative specification and generation of digital libraries. JCDL 2002: 263-272, Portland, Oregon (2002)
6. Novak, J.D., Learning, Creating, and Using Knowledge: Concept Maps as Facilitative Tools in Schools and Corporations. 1998, Mahwah, NJ: Lawrence Erlbaum.
7. Marshall, B., et al. Convergence of Knowledge Management and E-Learning: the GetSmart Experience. in JCDL. 2003. Houston.
8. Rajaraman, K. and A. Tan. Knowledge Discovery from Texts: A Concept Frame Graph Approach. in Eleventh International Conference on Information and Knowledge Management (CIKM'02). 2002. McLean, Virginia, USA.
9. Saad, A. and A.-R.M. Zaghoul. A Knowledge visualization tool for teaching and learning computer engineering knowledge, concepts, and skills. in ASEE/IEEE Frontiers in Education Conference. 2002. Boston, MA.
10. Saito, H., et al. A Semi-Automatic Construction Method of Concept Map Based on Dialog Contents. in Sixth International Conference on Computers in Education (ICCE2001). 2001. Seoul, South Korea.
11. Suleman, H.: Open Digital Libraries, Ph.D. dissertation, Virginia Tech, 2002
12. Zhu, Q.: 5Sgraph: A Visual Modeling Tool for Digital Libraries. Master Thesis. Virginia Tech. <http://scholar.lib.vt.edu/theses/available/etd-11272002-210531/> (2002)