SEMI-AUTOMATED MAP GENERATION FOR CONCEPT GAMING

Lauri Lahti and Jorma Tarhio
lauri lahti at hut fi, jorma tarhio at hut.fi
Helsinki University of Technology, Department of Computer Science and Engineering
P.O. Box 5400, FI-02015 HUT, Finland

ABSTRACT

Conventional learning games have often limited flexibility to address individual needs of a learner. The concept gaming approach provides a frame for handling conceptual structures that are defined by a concept map. A single concept map can be used to create many alternative games and these can be chosen so that personal learning goals can be taken well into account. However, the workload of creating new concept maps and sharing them effectively seems to easily hinder adoption of concept gaming. We now propose a new semi-automated map generation method for concept gaming. Due to fast increase in the open access knowledge available in the Web, the articles of the Wikipedia encyclopedia were chosen to serve as a source for concept map generation. Based on a given entry name the proposed method produces hierarchical concept maps that can be freely explored and modified. Variants of this approach could be successfully implemented in the wide range of educational tasks. In addition, ideas for further development of concept gaming are proposed.

KEYWORDS

concept map, computer game, map generation, open access

1. INTRODUCTION

Concept maps have proven to be a practical and illustrative way to represent relationship of interrelated concepts (Gaines and Shaw, 1995). Typically, concept maps are graphs of nodes labeled with concepts. Labeled directed edges depict the relationships of these concepts. Concept maps are actively used in education both as a collective and an individual tool for notation. However there has not been much initiative for using concept maps as an interactive educational tool. Instead, the concept maps have been typically used as a static illustration that is fixed in its structure and contents (Leake *et al.*, 2003).

Anyway, there have been some proposals for providing more challenging and personalized use for concept maps in computer-aided education. One of them is the idea of *concept gaming* that combines contextually bound concept map building with motivating factors of an interactive game (Eronen *et al.*, 2002). The aim of concept gaming is to provide a learning environment that offers motivation for conceptual analysis through challenge and excitement. Originally, concept gaming is developed for educational use in primary school. Concept gaming has been implemented so far in one prototype application.

In this paper we propose a new semi-automated map generation method for concept gaming using online content. This method is intended to be a supplement to the original concept gaming prototype and our group continues the work of Eronen et al. (2002). The original prototype consists of a fully functional game platform and a map editor. However, getting started with the prototype may be hindered by the need for constructing concept maps always manually in the first place. Thus it seemed worthwhile to develop further the idea of concept gaming to support semi-automated map generation. To make it easier to understand functioning of the new tool, it is necessary to explain some basic properties of the original prototype.

2. CONCEPT GAMING

Lahti, L., & Tarhio, J. (2008). Semi-automated map generation for concept gaming. Proc. IADIS International Conference Gaming 2008 (part of MCCSIS 2008): Design for Engaging Experience and Social Interaction, 25-27 July 2008, Amsterdam, the Netherlands (eds. Xiao, Y., & ten Thij, E.), IADIS Press, 36–43. ISBN 978-972-8924-63-8.

2.1 Overview

The basis of the concept gaming approach introduced by Eronen et al. (2002) is a set of *game schemes* that are run with certain concept maps. Before the gaming can start, a concept map must be drawn with an integrated editor or loaded from a file. Then in the actual gaming phase, some parts of the concept map are shuffled or hidden from the user and she is asked to complete it. So far, the prototype provides five different game schemes. For simplicity, the nodes in the map are often called as concepts and the edges as relations.

In the game scheme of removed concept names, the map is shown without concept names and the user is required to put them into correct place from a list. Figure 1 gives an example of this scheme. In the scheme of removed relation names it is then relation names that need to be added from a list respectively. In trivia scheme the user is provided with a relation and its starting concept and the user is required to select a correct ending concept from a list. The two other schemes are shuffled concepts and shuffled relations which both show the map and require arranging randomly shuffled concepts or relations respectively.

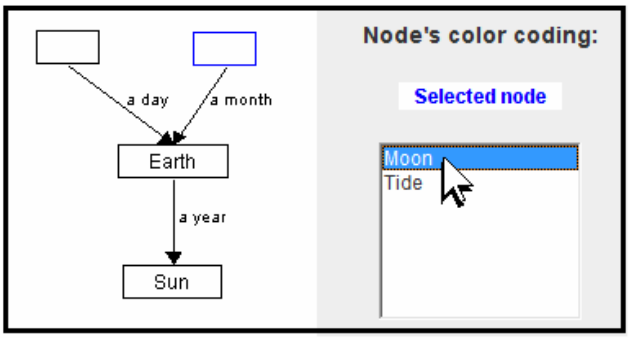

Figure 1. Game scheme of removed concept names in concept gaming prototype (a detail).

In all of these game schemes the performance of the user is evaluated based on correctness and speed. The user is provided with an instant feedback after each move and in the end the final score and the correct answer is shown. Besides these five game schemes it is possible to implement new schemes. Also the concept maps needed for gaming can be easily created, saved, and shared. Once a concept map has been created it can be applied to any one of the available game schemes.

Concept gaming differs from conventional learning games due to its generative nature. Conventional learning games are typically inherently fixed on a single subject whereas the concept gaming scheme can generate games on various different subjects depending on the concept map given as an input. Reflecting the findings of Chiu (2001), concept gaming allows to use a single concept map to create many games and these can be chosen so that personal needs of the user are taken flexibly into account. Concept gaming in its original form is intended to meet particularly well the educational needs of students in primary school. Through adjustable visualizations children can get valuable support for taking over the organization of concepts and their structural relations.

2.2 Impact of concept gaming

Concept gaming provides a framework for handling conceptual structures that are defined by a concept map and thus different kinds of concept maps can be created responding to the current learning goals. For example, the gaming can be based on maps created by a teacher, or students can create maps themselves. In the individual process of building a concept map, the personal mental models of a student can be addressed in a specific way.

The influence of concept gaming has been analyzed from student's and teacher's point of view. The original findings can be summarized as follows (Eronen *et al.*, 2002). Concept gaming seems to support using different learning skills and methods as well as to construct and visualize new inner cognitive structures. However, abstract conceptualized thinking skills are needed. Concept gaming brings excitement to learning experience. On the other hand, the learning space on the screen is limited and the fixed model

answers might suppress the student's own ideas. From a teacher's perspective, concept gaming seems to help in understanding and meeting the individual needs of the student. Concept gaming provides a practical way to measure the learning progress. In addition, once the concept maps have been prepared they can be easily shared and exchanged with anyone.

In later experiments it was noted that talented students felt that playing with maps constructed by the teacher hindered their own creativity and they wanted more challenge (Silander, 2006). Therefore we recommend that students should be allowed to construct maps also themselves according to their interests. Furthermore, the construction process of a concept map can provide useful information about problematic zones in learning. Thus the methodology of concept map scripts could be implemented to monitor construction history (Rautama *et al.*, 1997).

3. NEED FOR CONCEPT MAP RESOURCES

3.1 Challenge of sharing contents

Concept gaming was aimed to serve in a variety of real-life learning scenarios. For example in basic education there is a demand for easy applications that can be effectively used to represent new structural information and to share it instantly (Majid *et al.*, 2006). The benefit of concept gaming depends much on the number and quality of the ready-made concept maps available in the format of the application. Eronen et al. (2002) suggest that a web-based database could give additional value and decrease the workload required for starting using the application.

However there is no specification format of concept maps which is widely accepted. There exist several formats that do not share common notation or encoding. Typically a specification consists of a node list with spatial information and a relation list. Such a specification is straightforward to encode for example with XML. Some competing commercial formats include MindGenius (MindGenius, 2008), MindManager (MindManager, 2008), and MindMapper (MindMapper, 2008). There is development for producing open standards for concept maps going on as well. Open formats specific to concept maps include CMapTools (CMapTools, 2008), FreeMind (FreeMind, 2008), and Topic Maps (Topic Maps, 2008). However, it is characteristic to open format development that it may converge slowly. Thus stable and easy-to-use applications for inexperienced users are not yet available.

The absence of both common encoding and conversion tools makes sharing concept maps challenging especially if one wants to preserve the ability for further editing. Many tools are oriented to provide the final map as an image file, such as a jpeg image, which leads to losing structural encoding. The tools often allow saving the map also in an editable format but typically the format is not supported by other tools. There are some web portals dedicated to sharing concept maps but they as well often lack the possibility for allowing further editing (Cañas *et al.*, 2004). This kind of strategy severely restricts reuse of many maps and causes a great loss of efforts used in them. Also a major challenge is that apparently there is not any practical way to make efficient searches among concept maps. Finding a concept map matching sufficiently your requirements can easily take more time than creating one by yourself from the scratch.

3.2 Semi-automated concept map generation

While standardization of concept map encoding has not yet found a full consensus and competing non-compatible formats exist, it seems that some transitional methods are needed to cope with current demand of sharing structural knowledge. Regardless of the decision of format used for representing structural knowledge it is also important to develop tools for automated creation of structural knowledge (Sue *et al.*, 2004). It appeared that concept gaming due to its unconventional approach could address these issues to some extent.

It was considered that for generation of maps for concept gaming in ordinary educational context it could be reasonable to make some trade-off between the efficiency and the accuracy. There are some highly sophisticated applications for automated text analysis and making summaries but they are typically computationally intensive and require use of complex models. Furthermore, the automated language processing typically needs large corpus and extensive indexing to work well. Therefore there seems to be a

need for a simple approach that works sufficiently even with low computational resources that are present in many educational instances. Avoiding unnecessary complexity supports the principles of providing equal possibilities to attend educational material for all (Teo and Gay, 2006). In addition, it needs to be emphasized that concept gaming does not aim to provide static learning content or an unambiguous solution. Thus the proposed rather robust method for creation of concept maps should be seen as an unrestrained starting point for exploration that actively encourages the user to initiate further improvement.

Development of concept gaming presented in this paper was greatly inspired by the emergence of rather reliable online knowledge resources on the Web. Due to fast increase in the open access encyclopedias, tutorials and article databases it seemed feasible to use online content as a source for semi-automated concept map generation. For educational use this would guarantee having always up-to-date material available. We decided to utilize Wikipedia, the free encyclopedia, which is maintained and constantly extended by a global community of volunteers (Wikipedia, 2008). We used only the English version of Wikipedia that consisted of about 2.4 million articles in May 2008. Actually, anyone is allowed to add or modify contents of the articles in the Wikipedia. The project of Wikipedia relies on the idea that through a collective iterative refinement process each article should reach an acceptable level of detail and accuracy. Naturally, the modification of articles is open to abuse but all the changes are stored in logs and some tracking features have been set up to alert in case of vandalism.

4. CONCEPT MAP GENERATION FROM THE WIKIPEDIA

4.1 Structural knowledge in the Wikipedia

In our new *map generation tool* the generation of concept maps is based on the standardized structure of the Wikipedia articles. The knowledge is extracted from the web pages according to the html tags that specify the text segments represented in the article.

The knowledge content available in the Wikipedia can be classified and approached in many ways. The basic principles are described here in short. All the articles in the Wikipedia can be found by entering the entry name into a search box. In the case of two or more similar entry names, the article describing the most popular meaning is primarily offered. If other meanings exist, it is indicated in the beginning of the article and they can be reached through special disambiguation pages.

Typically all articles have been tagged to belong to one or several categories. There are specific pages to show all the entries belonging to a certain category and the hierarchy of the categories. For a large number of articles there is a special section of table of contents in the beginning of it. The table of contents typically lists the headings used in the article and provides hyperlinks to these sections. According to the documentation of the Wikipedia, for each page with more than three headings, a table of contents is normally automatically generated from the section headings. This is avoided only if a user has set a preference to turn it off (Wikipedia documentation, 2008).

There does not seem to be a systematic way of writing headings for articles and due to this the headings vary a lot between different articles. Despite of this ambiguity, it was considered that when observing the whole entity of the Wikipedia articles, on average the table of contents is the most condensed and uniform representation about the structure of each article. Furthermore, it was evaluated that with quite high probability the table of contents seems to contain some specific keywords and perspectives describing the subject of the article. Especially with such longer articles that are divided into subsections on several levels it appears that table of contents can provide some simple knowledge about the hierarchy of the specific keywords and perspectives describing the subject.

Since the articles in the Wikipedia are written on voluntary basis, the topics are not covered in a systematic order of importance. Some entry names are still missing an article although they would be surely mentioned in a traditional encyclopedia. There are also many articles that have been barely started. Among one hundred randomly chosen articles we measured that only 34 % were long enough to be supplied with a table of contents and of these 47 % had more than one level of hierarchy in it. However, among many educationally valuable articles there is much more often a multi-level table of contents available. We took one hundred main class names from Universal Decimal Classification system (Universal Decimal Classification system, 2008) used in libraries and evaluated corresponding articles in the Wikipedia. 66 % of

these articles provided a table of contents having at least two levels of hierarchy and of these 82 % had second-level headings under more that one first-level heading. This all confirmed us that among educationally important Wikipedia articles the table of contents can provide sufficiently hierarchical knowledge to generate maps for satisfactory concept gaming.

4.2 Process of concept map generation

Both the original concept gaming prototype and the map generation tool have been written in Java and they can be run on common operating systems with a graphical user interface. When a teacher wants to generate new concept maps, the map generation tool requires her to define the entry name that is wanted to be as the topic of the map. Then the application establishes a web connection to the site of Wikipedia using a web address extended with the entry name (for example, http://www.wikipedia.org/Horse). If an article matching to the entry name is found its table of contents is extracted according to the html tagging. Then the application forms a concept map object following the specifications of the concept gaming prototype and saves it in a file. The generated map consists of a hierarchical representation of nodes including the headings taken from the table of contents. The edges between the nodes are directed from the root to the leaves.

For simplicity, the application produces the hierarchical tree structure with a condensed layout. Even if only a small amount of nodes is retrieved, getting them automatically is enough to make a difference. Just a few tentative nodes showing some central subordinate concepts about the topic can give valuable inspiration and support building a comprehensive map. The teacher is encouraged to make some finishing for this initial concept map according to the context-specific needs. For example, the teacher can add or remove nodes and edges, or reorganize them. Figure 2 illustrates a map generated from the Wikipedia article having entry name “Horse” in May 2008.

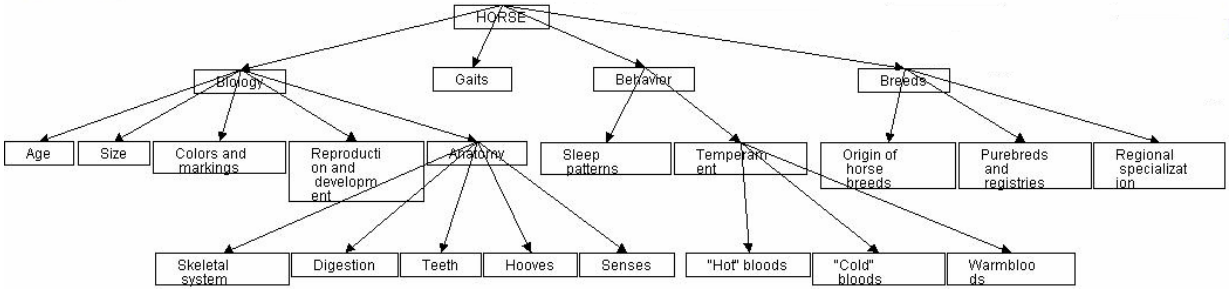

Figure 2. Concept map generated from the Wikipedia article about “Horse”.

The generated maps can be delivered to the concept gaming prototype and used in all available game schemes discussed in Chapter 2.1. If the teacher does not add edge names manually the maps can currently be used meaningfully only in the game scheme of removed concept names (see Figure 3). If the concept map has a hierarchical structure and edge names are not present the concept map reduces to a *mind map*. Thus the maps generated by the application are initially mind maps which can be easily expanded and improved by the teacher. Naturally, the map generation tool can be used by a student as well. We think that the initial structure created by the application provides a flexible starting point for associations and fosters the creativity for further map building in a pedagogically generous way.

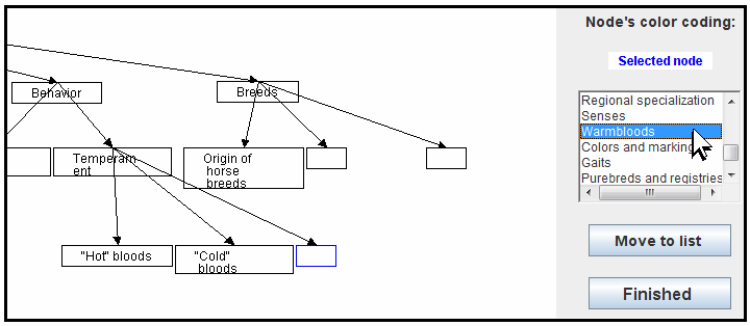

Figure 3. Game scheme of removed concept names that is based on concept map about "Horse" (a detail).

4.3 Extensions to map generation

The semi-automated map generation gives a methodology to produce content for learning process but it is the user community who finally makes it to flourish. The map generation should be considered as just the first step to the diversity of activities based on the potential of concept gaming.

Concept gaming can provide games that aim at enhancing students to conceptualize the real world and features of everyday life. Instead of a separate virtual gaming experience, we want to encourage people to see challenging gaming opportunities everywhere related to building personal conceptual understanding of all aspects of life. We think that it is valuable to motivate people not only to make abstractions with concepts but also to establish their own definitions and attitudes towards them through gaming process. Possibly, in the long run, the personal conceptualization produced in gaming could be gathered as a diverse, collective resource for better mutual understanding. Same kind of approach has been adopted for example in the Conzilla project for concept browsing (Naeve, 2001).

In educational context the concept gaming has originally aimed at level of primary school but we try to broaden the use for other groups as well. The textual information of nodes and edges could be replaced with images or animations and accompanied with auditory cues. Thus enhanced concept gaming could offer tailored training for kindergarten and children with special needs. We have also considered expanding the principles and applicability of concept gaming from school environment to other professional working fields. Note, that for example rebuilding a map from shuffled concepts is related to a certain brainstorming technique and offers inspiration for creative problem solving.

5. FURTHER DEVELOPMENT

We will study other methods for gathering online knowledge for concept map generation. There are several ways to approach the contents of Wikipedia by simply downloading required html files. It would be possible to identify all entry names mentioned in an article and to form a hierarchy based on the network of these hyperlinks. An application could even crawl through these links to related articles and establish more complex map structures than trees. Furthermore, scanning all contents of these related articles could reveal even more detailed structural relationships. One could also use information about the subject categories into which each article is tagged to. Finding the most related concepts in an article could be based on matching with the same subject categories. Besides the Wikipedia encyclopedia, one could extract simple structural knowledge from various other web sites as well. For general and multidisciplinary topics the knowledge might be gathered, for example, from subject directories of Yahoo and Google or library classification vocabularies. For more specific topics, it might be worth scanning through some popular portals and tutorials in the field. Besides map generation, we will also develop the original concept gaming prototype further. In all phases of concept gaming the application should keep track of user activities in order to provide an adaptive response and to analyze learning profiles.

In general, there seems to be too many competing and non-compatible formats for representing concept maps which may seriously limit adoption and collaboration in using them in the computer-aided way. There is potential in developing further distant collaboration and easy integration of concept maps with other

material on the Web (Rautama and Tarhio, 1998). Also developing sophisticated and universally disposable methods for concept maps suffers from the lack of common notation. Since representing a concept map actually requires only very simple data structures we are planning to make new conversion tools from several concept map formats to the format of our system. It still needs to be emphasized that the principles we have introduced for concept gaming will stay applicable with any chosen standard.

To address the need for processing conceptual information on deeper semantic level we are planning to incorporate ontological knowledge to concept gaming. The linked open data movement, largely supported by World Wide Web Consortium, has been assigned as a major force in developing current web resources to a semantic web (Linking Open Data project, 2008). For example, the DBpedia project (DBpedia project, 2008) licensed under the free GNU license provides a flexible data model for representing semantic information extracted from the Wikipedia and allows making online queries of its contents. By February 2008, the community effort of the DBpedia had produced descriptions of 2.1 million topics using Resource Description Framework (Resource Description Framework, 2008) that allows describing systematically various properties of a topic, such as a label, a short and a long abstract, links to corresponding Wikipedia article and image, and different language versions.

In addition, the DBpedia provides classification of topics based on categories of the Wikipedia and the closely related Yago classification (Suchanek *et al.*, 2007) and WordNet lexical thesaurus service (Fellbaum, 1998). Also the external links and geo-coordinates concerning the topics can be taken into account. We think that openness with data combined with harmonization of open standards can revolutionize the paradigms of collaborative learning. The semantic web based on the Resource Description Framework and its extension the Web Ontology Language (Horrocks *et al.*, 2005) should enable people to navigate from one data source to another related one following the ontologies defined for them.

Concept gaming can be seen as a way to support both intellectual competence and creating collaborative consensus over ontologies of the semantic web. Emphasizing gaming aspects for the community of volunteers can help to ensure inspiration and challenge needed for continuous collective improving of present ontologies. Like with the Wikipedia, building the ontologies for the semantic web on voluntary basis does not guarantee systematic and detailed coverage of all areas of human knowledge. Here concept gaming could be used to attract people to contribute especially in topics that do not belong to the mainstream and would otherwise be covered slowly. Also, to define collectively the most controversial parts of ontologies one could benefit from game-driven approach in a similar fashion as the Wikipedia reaches consensus that is based on the most supported and viable opinions.

With concept gaming the community could even naturally assist in supervised machine learning techniques. Google Image Labeler (Google Image Labeler, 2008) is a working example of gaming concept that harnesses volunteers to help to form semantic classification rules. However, it is typically difficult to evaluate reliably the quality of produced content and reward the volunteers accordingly in a motivating way. Since building of ontologies is based on critical reasoning that increases collective understanding it seems that concept gaming can inherently provide features that support high quality and motivation. We will further develop schemes of concept gaming that can assist in building ontologies for the semantic web.

6. CONCLUDING REMARKS

Concept gaming is a fascinating tool for computer-aided education. We implemented a semi-automated map generation tool that broadens its use possibilities. Moreover, we presented ideas for further development of concept gaming. Our application generates concept maps based on articles of the Wikipedia online encyclopedia. The created maps can be further edited if wanted. The maps can be imported to the original concept gaming prototype and applied in a variety of game schemes. The new tool aims to ease the burden of creating new concept maps for illustrating hierarchical relations for educational purpose. In concept gaming, even a single concept map can offer numerous learning approaches since both teachers and students can explore freely with structural configurations. Different kinds of concept maps and game schemes can be created responding to the individual learning goals. The new tool widens the usability of original concept gaming allowing flexible creation of concept maps based on up-to-date online content. In the current stage, the application extracts structural knowledge from the table of contents of a Wikipedia article and produces a hierarchical tree of it.

REFERENCES

Cañas, A. et al., 2004. CmapTools, Web Pages & Websites. Technical Report IHMC CmapTools 2004-01. Institute for Human & Machine Cognition. Pensacola, Florida, United States. http://cmap.ihmc.us/Publications/WhitePapers/CmapTools,%20Web%20Pages%20&%20Web%20Sites.pdf [June 2008]

Chiu, C., 2001. The authority structure problem of computer supported collaborative concept mapping system for elementary students. *Proceedings of IEEE International Conference on Advanced Learning Technologies,* Madison, Wisconsin, United States, pp. 57 – 60.

CMapTools, 2008. http://cmap.ihmc.us [June 2008]

DBpedia project (2008). http://wiki.dbpedia.org [June 2008]

Eronen, P. et al., 2002. Concept Gaming. *Proceedings of the Tenth International Conference on Computers in Education*, Los Alamitos, California, United States, pp. 997 – 1001.

Fellbaum, C. (ed.), 1998. WordNet – An Electronic Lexical Database. MIT Press, Cambridge, Massachusetts, United States. http://wordnet.princeton.edu [June 2008]

FreeMind, 2008. http://freemind.sourceforge.net/wiki/index.php/Main_Page [June 2008]

Gaines, B., and Shaw, M., 1995. Collaboration through concept maps. *Proceedings of the First International Conference on Computer Support for Collaborative Learning*, Bloomington, Indiana, United States, pp. 135 – 138.

Google Image Labeler, 2008. http://images.google.com/imagelabeler/ [June 2008]

Horrocks, I. et al., 2005. OWL Rules: A Proposal and Prototype Implementation. *Journal of Web Semantics,* Vol. 3, No. 1, pp. 23 – 40.

Leake, D. et al., 2003. Aiding knowledge capture by searching for extensions of knowledge models. *Proceedings of the Second international conference on Knowledge capture.* Sunibel, Florida, United States, pp. 44 – 53.

Linking Open Data project, 2008. World Wide Web Consortium, Semantic Web Education and Outreach Interest Group Community Project. http://esw.w3.org/topic/SweoIG/TaskForces/CommunityProjects/LinkingOpenData [June 2008]

Majid, O. et al., 2006. The Effect of Web-Based Concept Mapping in Learning Enhancement. *Proceedings of the Sixth IEEE International Conference on Advanced Learning Technologies*, Kerkrade, Netherlands, pp. 954 – 958.

MindGenius, 2008. http://www.mindgenius.com [June 2008]

MindManager, 2008. http://www.mindjet.com/products/mindmanager_pro/ [June 2008]

MindMapper, 2008. http://www.mindmapper.com [May 2008]

Naeve, A., 2001. The Concept Browser – a new form of Knowledge Management Tool. *Proceedings of the Second European Web-based Learning Environments Conference*, Lund, Sweden, pp. 151 – 161.

Rautama, E. and Tarhio, J., 1998. Sharing concept maps on the Web. *Proceedings of the Sixth International Conference on Computers in Education*, Beijing, China, pp. 273 – 280.

Rautama, E. et al., 1997. Supporting Learning Process With Concept Map Scripts. *Journal of Interactive Learning Research*, Vol. 8, No. 3/4, pp. 407 – 420.

Resource Description Framework, 2008. RDF Primer, World Wide Web Consortium Recommendation, 10 February 2004. http://www.w3.org/TR/rdf-primer/ [June 2008]

Silander, P., 2006. Private communication.

Suchanek, F. et al., 2007. YAGO – A core of semantic knowledge unifying WordNet and Wikipedia. *Proceedings of the 16th international World Wide Web conference*, Banff, Alberta, Canada, pp. 697 – 706.

Sue, P. et al., 2004. A new approach for constructing the concept map. *Proceedings of IEEE International Conference on Advanced Learning Technologies,* Joensuu, Finland, pp. 76 – 80.

Teo, C. and Gay, R., 2006. A knowledge-driven model to personalize e-learning. *Journal on Educational Resources in Computing*, Vol. 6, No. 1.

Topic Maps, 2008. http://www.topicmaps.org [June 2008]

Universal Decimal Classification system, 2008. Summary of classification. UDC Consortium. Hague, Netherlands. http://www.udcc.org/scheme.htm. [May 2008]

Wikipedia, 2008. The free encyclopedia. http://www.wikipedia.org [June 2008]

Wikipedia documentation, 2008. Help section, Table of contents (TOC). http://en.wikipedia.org/wiki/Wikipedia:Section#Table_of_contents_.28TOC.29 [June 2008]