Knowledge exploration with concept association techniques

Chen-Chung Liu and Shih-Hsun Fan Chiang
Graduate Institute of Network Learning Technology,
National Central University, Jhongli City, Taiwan, Republic of China

Chih-Yueh Chou
Department of Computer Science and Engineering, Yuan-Ze University,
Jhongli City, Taiwan, Republic of China, and

Sherry Y. Chen
Graduate Institute of Network Learning Technology,
National Central University, Jhongli City, Taiwan, Republic of China

Abstract

Purpose – Exploratory learning is regarded as an important ability for developing knowledge from open environments. During the exploration, learners not only need to acquire new information based on their current interests, but also they need to form new perspectives by incorporating new knowledge into their previous knowledge. This paper seeks to address these issues.

Design/methodology/approach – To this end, this paper proposes an approach that includes a concept association bank to recommend related concepts in a domain based on the goal of an exploration. By doing so, learners’ knowledge can be expanded beyond their current understanding. An experiment was conducted to investigate how the proposed approach facilitated the learners’ exploration.

Findings – The results indicated that the concept association bank is a useful mechanism to help learners gain new understanding, including providing exploration directions, reducing complexity and cognitive load, facilitating data- and goal-driven exploration strategies, and commenting on new understanding. The implications of these results are discussed.

Originality/value – Current recommendation systems emphasise a data-driven strategy, which seeks isolated pieces of information, instead of suggesting directions related to their exploration goal. The problem with such an approach is that learners’ exploration will be limited by their existing knowledge. Thus, this paper presents an approach to support both data- and goal-driven strategies.

Keywords Worldwide web, Learning, Knowledge mining

Paper type Research paper

Introduction

Exploratory learning (Jonassen and Mandl, 1990; Rieman, 1996) is regarded as an important ability for developing knowledge from open environments, such as the web and internet databases, since they contain valuable and comprehensive information for learners. Learners need to uncover interrelated concepts and topics to explore a domain in an exploratory learning environment. More specifically, this is a self-initiated and goal-oriented process, during which learners independently explore information in databases and gradually acquire new knowledge to extend their understanding.

The authors appreciate the work of their anonymous reviewers in suggesting revisions for this paper. This research was partially funded by the National Science Foundation under 98-2511-S-008-004-MY3, 98-2631-S-008-001, and 96-2524-S-008-002.
However, learners may easily experience disorientation and find the cognitive load too heavy (Conklin, 1989; Madrid et al., 2009) due to the huge amount of information. Therefore, facilitating information extraction and assisting learners to explore the environments become critical issues to support exploratory learning. In other words, there is a need to develop a mechanism that can help students to investigate the domain.

To address this issue, a number of recommendation mechanisms have been devised, such as content-based filtering (Mooney and Roy, 2000) and collaborative filtering (Good et al., 1999). Many of these mechanisms recommend information in order to inform the users about new items such as books, etc. However, learners not only need to acquire new information based on their current interests; they also need to form new perspectives by incorporating new knowledge into their previous knowledge (Siemens, 2005). While new knowledge is built on previous knowledge, the development of new knowledge usually consists of associations that had not been thought of before (Shneiderman, 2000). Therefore, it is necessary to provide learners with recommendations for further expanding their current understanding. However, the information objects recommended by current recommendation systems are “overspecialised” (Adomavicius et al., 2005) and thus learners’ exploration of the recommendation will be limited by their existing knowledge. This is due to the fact that the current recommendation systems emphasise a data-driven strategy, which seeks isolated pieces of information based on their current interests, instead of suggesting new directions which would help learners to expand their current understanding.

Expanding current understanding of a domain can be achieved by associations, which indicate correlations between learning objects (Ouyang and Zhu, 2008). In order to facilitate learners’ associations this paper developed a mind map tool equipped with association support to provide learners with concepts related to their current understanding. The mind map tool provides a kind of association support that recommends concept keywords as new directions for exploration. The association support facilitates learners to explore a domain using the work of experts in a particular community. The concept associations were not provided by teachers or stakeholders; instead they were obtained by mining academic articles contributed by a majority of experts in a professional community. Recently, online academic collective databases have become an important channel for accumulating the work of experts in a community. It is hoped that the concept associations obtained from online academic collective databases will help to stimulate divergent exploration directions and deepen the exploration of a domain. However, the strategies that learners apply to explore a domain may vary (Hill and Hannafin, 1997). Therefore, it is necessary to analyse their exploration strategies when they are provided with the mind map tool. To this end, we conducted an empirical study to examine how the concept association approach affects exploratory learning. The ultimate goal of the empirical study was to develop a framework based on the answers to the following research questions:

*RQ1.* How can the concept association support help learners to identify exploratory directions?

*RQ2.* How can the concept association support reduce the learners’ cognitive load?

*RQ3.* How can the concept association support affect learners’ information seeking strategies?
Related works
Exploratory learning involves a learner iteratively seeking information from open environments to advance their understanding of a self-generated topic. People usually exhibit such exploratory learning behaviours when they find their current state of knowledge is inadequate to achieve their goals. This suggests that people engaging in exploratory learning do not really know what might be useful for them, and therefore may not be able to specify the salient characteristics of potentially useful information objects (Belkin, 2000). In addition, the open learning environments, such as the web and databases, are not organised in a clear structure (like that of textbooks) that can guide learners to independently explore information through topics to extend their understanding. As a result, a recommendation system is critical to provide not only guidance in the form of navigational support but also facilities to help learners learn new concepts through information seeking (Hubscher and Puntambekar, 2001). Previous studies have proposed various approaches to provide recommendations to help learners explore a domain, which can be divided into content-based and collaborative filtering approaches as detailed below.

Content-based recommendation systems
The content-based approach provides recommendations to users primarily based on information retrieval (Baeza-Yates and Ribeiro-Neto, 1999) and information filtering (Belkin and Croft, 1992). This approach calculates the fitness of information objects to users’ profiles in order to recommend information objects that fit the current interests of the users. It thus involves two main tasks to provide recommendations: extracting the characteristics of the information objects and building the user profiles. Regarding the task of extracting characteristics of information objects, the recommendation systems have to process the information objects to obtain the computational features of these objects. The recommendation systems also have to trace users’ exploration behaviours to build computational user profiles that reflect the interests of the users. The recommendation systems can determine whether an information object should be recommended to a user by calculating the similarity between the information object and the user’s interests.

Previous studies have proposed different content-based approaches to perform the above two tasks in order to recommend information objects to users. Most of the proposed approaches utilise keywords as the content information to calculate the features of information objects. For example, keywords of books (Woodruff et al., 2000), web sites (Balabanovic and Shoham, 1997), news (Billsus et al., 2000), and professional academic papers (Torres et al., 2004) were extracted from the information objects to represent the features of each information object by the use of information retrieval techniques such as term frequency – inverse document frequency (TF-IDF; Salton et al., 1975). The recommendation systems trace users’ information seeking behaviours to locate the information objects of interest to the users. By processing the information objects which the users have recently retrieved, user profiles can be constructed to represent users’ interests. For instance, the recommendation systems proposed in prior studies (Joachims et al., 1997; Hsu, 2008) trace a user’s web information seeking behaviour to determine the user’s interests. The content-based recommendation systems may then compute whether an information object should be recommended to the user according to the user’s profile and the features of the information object.
Such content-based approaches can recommend suitable information objects based on users’ past interests. However, the information objects that are not reflected in their interests would not be recommended to them. Thus, the users might be restricted to those objects that are very similar to the ones they already know about (Tang, 2008). This is why, the information objects recommended by this approach are regarded as “overspecialised” (Adomavicius et al., 2005) or trapped in a “similarity hole” (McNee et al., 2006). Such an overspecialised recommendation or similarity hole creates an obstacle to exploratory learning. This is due to the fact that the learners may not be aware of the most useful objects so the information objects that they have already retrieved may not sufficiently reflect their interests. Thus, there is a need to recommend the information objects outside their past information seeking experiences so that their current understanding can be increased.

**Recommendations based on collaborative filtering**

The collaborative filtering approach is one of the solutions to the overspecialisation and similarity hole issue with content-based recommendations. Unlike the content-based recommendation which recommends information objects based on the past interests of a single user, collaborative filtering makes recommendations on the basis of a community of users who share similar interests. The approach has been extensively applied in e-commerce to help users locate items of interest. A frequently used approach is market basket analysis (Bigus, 1996), which refers to the application of data analysis techniques to databases that store transactions from consumers buying selections of different products. The aim of the analysis is to understand the association structure between the sales of the different products available. Once the associations are found, they may help in developing marketing policies. For instance, if there is a relationship between two products over time, then retailers can use this information to contact the customer, decreasing the chance that the customer will purchase the product from a competitor. This is a typical association rules approach which has been applied in many areas of e-commerce, including movie reviews (Miller et al., 2003), online food stores (Svensson et al., 2000), music suggestions (Chen and Chen, 2001), and online bookstores such as Amazon.com (Linden et al., 2003), etc.

This approach is also useful for exploratory learning since it can provide learners with information derived from the experiences of other learners. More specifically, an information object may be more helpful to a learner when another learner with a similar background considers it helpful. This approach generates an association between information objects that is valid among a certain group of learners within one specific domain to predict the rating value of the information objects. Based on the association, it is possible to recommend to a learner an object that other learners are interested in. For instance, if we find that learners are generally interested in X and Y at the same time, we would recommend Y to a learner who retrieves X.

**Recommendations based on goal-driven strategies**

Both the content-based and collaborative filtering approaches emphasise seeking isolated pieces of information that fit the current needs of a learner or a group of learners, which is a data-driven strategy. In addition to a data-driven strategy, exploratory learning in an open learning environment also requires a goal-driven strategy to make principled judgments related to the exploration goal (Land and Greene, 2000).
The difference between data- and goal-driven strategies is that learners use the former to locate each piece of related information and examine the details of each piece but they use the latter to integrate and link the related information to form a concept that relates to the goal of an exploration. Therefore, goal-driven exploration support not only helps learners to look for isolated resources but also suggests new directions that the learners have not yet noticed for further exploration.

To this end, the study presented in this paper proposes an approach, which addresses the association between the goal and the concepts presented in isolated information objects, instead of focusing on the association between information objects. In other words, the novelty of this approach lies within the fact that not only will isolated information objects be recommended, but we will also consider the goal of an exploration and then suggest the concepts related to the goal as advised by a community of experts.

**Conceptual association support**

This study proposes a recommendation system which facilitates learners to explore a domain by the use of both data- and goal-driven strategies. The mechanism, as shown in Figure 1, includes a mind map tool and a concept association bank. The data-driven strategy in this system is a mind map tool, with which isolated search results on the web can be integrated in a graphical representation. In other words, the mind map in this mechanism serves as a knowledge integration platform that helps learners to link, organise, and reflect upon their search results on the web. The goal-driven strategy in this system is a concept association bank to recommend not only isolated papers, but also related concepts in a domain, based on the goal of the learner’s exploration. The concept association bank is generated by analysing the keywords and keyword associations in the research collections of a professional community. In other words, the concept associations in the bank reflect the concepts and associations that the experts in the community consider important. It is hoped that the recommendation of concepts based on the concept association bank can improve learners’ conceptual awareness of the domain and thus facilitate them to achieve their exploration goal. The detailed descriptions of these elements are given below.

**Concept associations**

This research developed a concept association bank to facilitate goal-driven exploration. The concept association bank was extracted from a collection of papers using data mining techniques because such techniques would help to process a large volume of text material in order to identify expert knowledge. Such a bank can provide recommendations for exploring new directions. Among various data mining techniques, TF-IDF and the association rule mining technique (Agrawal et al., 1993) were applied to mine concept association in the community because the two techniques are useful for extracting important concepts and essential association between these concepts embedded in the paper collections of a professional community. The method of applying data mining techniques to support goal-driven strategies involves the following computational tasks:

- selecting the paper collections and the categories in which documents related to the domain are to be explored;
- collecting the title, author(s), keywords, and the abstract of the documents in the selected categories;
extracting the keywords (i.e. the concepts that need to be explored) of the selected papers with TF-IDF;

discovering the concept association by using the association rule technique; and

applying the discovered concept associations to recommend concepts for further exploration.
The aforementioned tasks suggest that the association rule plays an important role in producing the concept association bank. This rule is extensively applied to discover the co-presence of commercial items in transaction databases for marketing purposes. This study discovered the co-presence of concepts (i.e. items in the association rule technique) in documents (i.e. transactions in the association mining technique) to uncover the concept association rules in paper collections of a domain.

Concept association rules are keyword sets that are frequently co-present in papers. The support of a concept association rule, that is, in how many documents two keywords are co-present, was calculated to represent the levels of commonness of these associations. Table I displays an example of papers with different sets of keywords. In this example, the association rule mining technique will discover several concept association rules, including 3-support rule \(<A, C>\), 2-support rule \(<C, E>\), and many 1-support rules like \(<A, B>\) and \(<B, C>\). The support of the concept association rules reflects the popularity of these associations. For example, the association between concepts A and C commonly exists in the domain community while the association between concepts B and C was found in relatively few works. More specifically, three works associate concepts A and C while only one work associates concepts B and C with each other.

**Mind map tool**

Knowledge exploration on the web involves complex mental intellectual activities, such as linking and searching activities, which require a sophisticated representation to display complex relationships between various resources (Hubscher and Puntambekar, 2001). Since mind maps are widely applied to represent, integrate, and assess knowledge (Buzan, 1994; Liu et al., 2005), this study used mind maps to help students integrate their mental understanding and the data items they found on the web. More specifically the mind map, which was represented as a graphical knowledge integration platform, enables students to perform the conceptual association activities while exploring a knowledge domain. The mind map consists of the following elements:

- **Concept nodes.** Learners draw nodes to represent the important concepts in order to reflect their current understanding of a knowledge domain.
- **File nodes.** Learners add documents, e.g. paper files, presentation slides, and webpages.
- **Comment nodes.** Learners include a text node to comment on a document or webpage.
- **Author nodes.** Learners include an important author in the knowledge domain.
- **Links.** Learners link nodes of the above types that they consider to be related to each other.

<table>
<thead>
<tr>
<th>Document ID (transaction)</th>
<th>Concept keywords (items)</th>
</tr>
</thead>
<tbody>
<tr>
<td>P1</td>
<td>A, B, C</td>
</tr>
<tr>
<td>P2</td>
<td>A, C, E</td>
</tr>
<tr>
<td>P3</td>
<td>A, C</td>
</tr>
<tr>
<td>P4</td>
<td>C, E, F</td>
</tr>
</tbody>
</table>

Table I.

An example of using the association rule model for calculating concept association
Figure 2 shows a mind map as an example of how a learner explores knowledge on the web. This learner (1) identified five concepts which indicated his initial understanding of an exploration topic and were displayed as white ovals. From the nodes of the concept type, the learner (2) explored the web and added the results that he considered relevant to the exploration topic as nodes with the earth icon. The learner then (3) added the PDF documents that he found on the web or he already had in his computer into the mind map. The learner (4) gradually expands his understanding of the exploration topic by including and linking new concept keywords and documents that were considered related to the knowledge domain. In summary, the mind map tool was expected to facilitate learners to explore the knowledge domains at a data-driven level at which individuals locate resources on the web and their comments on the result can be integrated into the mind map.

The mind map includes the concept association bank, which helps learners initiate new exploration directions at a goal driven level by providing the following recommendations:

- **Concept associations.** The concept association rules suggest the associated concepts of a concept node in learners’ mind maps.
- **The list of related papers.** When suggesting a concept association to learners, the mind map provides a list of papers that contain the associated concepts as keywords.
- **Expert associations.** The concept association rules suggest a list of experts who wrote the documents containing the associated concepts.
- **Associations between concepts.** The association rules help learners to link a group of concepts by providing documents that contain all these concepts as keywords.

Figure 2 also shows how the concept association bank recommended concepts that were related to those concepts identified by the learner. The green ovals represent the concepts that were recommended by the concept association bank while the white ovals were those concepts identified by the learner himself. In the mind map, the learner first identified a concept by himself:

- From the self-identified keywords, the learner extended his exploration by requesting concept association recommendations.
- The concept association bank will recommend all the concepts that are related to the concept specified by the learner. The learner can then decide whether to accept the recommendations. If the learner decides to accept the recommendations, the recommended concepts will be added and linked to the concepts specified by the learner.
- In addition, the concept association bank will also recommend a list of related papers that are associated with the learner-specified concept and the recommended concepts.
- The learner expanded the exploration so that more and more concepts and resources were included in the mind map.

For instance, the mind map revealed that the learner’s exploration migrated from a general concept of “estimation” to more specific concepts such as “Bayesian estimation”
Figure 2.
An example of how a learner explores knowledge with a mind map and concept association.
and “Markov chain Monte Carlo” and included the search results related to these concepts in his mind map. Therefore, with the support of concept association learners cannot only receive recommendations of individual documents but also related important concepts for further exploration.

Method
Participants and procedures
An experiment was conducted to investigate how the proposed concept association bank assisted the learners, who were working on a specific domain, to explore a knowledge domain on the web. The experiment was conducted in a controlled setting because this study aimed to explore learners’ knowledge exploration behaviours, and thus several variables, such as the exploration tools and the time available, had to be controlled. The participants in this study – 25 students selected from a university – were working on theses related to the domain of Technology-enhanced Learning (TeL). A slight majority (14) were male. Most (20) of the participants were in the second year of a masters degree while five participants were doctoral students. As these students were studying TeL, exploring TeL domain knowledge on the web was one of the important tasks for their studies.

The participants were asked to freely explore their own thesis topics with the mind map, instead of being given specific topics. The exploration activity consisted of two sessions in two weeks. In each session, the participants spent 1 hour exploring the web with the mind map tool. During the exploration they needed to integrate the concepts they used to search the web and the resources they found in their mind maps. Their exploration behaviour was stored in a log file to identify the effects of the concept association support on domain knowledge exploration. We also conducted interviews afterwards to examine their perceptions in order to gain a better understanding of how the conceptual association support influenced their understanding of the topics. During the interviews, the participants were asked to review their mind maps and nominate important concepts for their topics out of all the concepts in their mind maps for 45 days after the second session of exploration activity.

The database and concept association rules
This study selected the ISI Web of Knowledge as the main database to construct the concept association bank because it is an important and widely used collection of documents for diverse academic communities. In addition, the database contains clear categories of disciplines by which documents can be easily classified to support the exploration of a specific domain in an academic community. Because the topics explored by the participants mainly focused on TeL, the study only used a portion of the database as the material for constructing the concept association bank. A total of 50,786 papers from 118 journals in the categories of educational research and development, special education, and ergonomics were selected. These papers contained abstracts and/or keywords provided by the authors. The papers that contained neither keywords nor abstracts were excluded.

The titles, authors, author-determined keywords, and abstracts of the selected papers were analysed to discover the concept association rules. The authors used a total of 71,519 keywords to describe their papers. However, 34,974 papers did not contain any author-determined keywords. The TF-IDF technique was therefore applied to extract

Concept association techniques

795
keywords from the abstracts for the papers that did not have author-determined keywords. A total of 152,823 keywords were extracted from the abstracts of those papers. Both the author-determined keywords and the keywords extracted by the TF-IDF technique were then analysed with the association rule technique to discover concept associations. As described above, the concept associations were produced on the basis of the co-occurrence of keywords. In total, 1,016,713 concept association rules were produced from the papers. These concept association rules were used as a knowledge base to provide concept association recommendations.

Data analyses

Analysis of learners’ mind maps. It was hypothesised that the concept association provided by the association rule techniques would improve learners’ understanding of their exploration topics. Therefore, the mind maps developed by participants were analysed to confirm the richness of students’ exploration efforts to understand the self-generated topics. If it is true that the initial level of understanding greatly influences learning behaviour (Song, 2003), learners’ self-determined keywords may play an important role in exploration. Consequently, both the keywords determined by students and those recommended by the concept association bank were analysed to confirm the effects of the conceptual association support. More specifically, the percentage of the keywords that were recommended by the concept association bank in the mind maps was calculated to reflect the changes of understanding caused by the concept recommendations. In addition, the percentage of recommended concepts in the list of important concepts that the learners identified in the interview was measured to confirm the effects of the conceptual association support on the enhancement of learners’ understanding.

Analysis of learner exploration activity. As the domain knowledge exploration involved in this study was primarily performed on a mind map tool, learners could perform rich mind mapping activities during their exploration. Consequently, the mind mapping activities were videotaped and observed to examine how the concept association bank affected learners’ exploration activities. The learners’ exploration patterns were analysed on the basis of the frequency of all nine main categories of activities during mind mapping: self-determining concepts, modifying concepts, deleting concepts, requesting concept recommendations, linking concepts, requesting associations between concepts, deleting links between concepts, browsing the web, and retrieving related papers. This study also conducted a sequential pattern analysis to reveal the learners’ exploration patterns with the concept association bank because this type of analysis has been applied in diverse areas such as communication patterns (Yamauchi et al., 2000) and web traverses (Ezeife and Lu, 2005) to understand users’ behaviour patterns. More specifically, this study used a transition diagram to conduct the sequential pattern analysis to reveal how the conceptual association activities, i.e. requesting concept recommendations and retrieving related papers, influence other exploration activities.

The series of activities performed by learners were coded into sequences according to the order in which activities were performed. The transition probability from activity \( a \) to activity \( b (a \rightarrow b) \) was calculated based on the percentage of transition \( a \rightarrow b \) among all transitions to \( b \). The probability of the transition \( a \rightarrow b \) reflects the tendency to perform activity \( b \) after activity \( a \). The transition probability was represented as the value attached to an arc between two states in the transition diagram.
**Results**

*Mind map analysis*

Table II presents the distribution of different nodes in the mind maps developed by learners. An average of 20.4 concept nodes appeared in the mind map of each learner. Among these 20.4 concept nodes, 9.88 were created by the learner, while 10.52 associated concepts were recommended by the concept association bank. In other words, 51.6 percent of concept nodes were created by the conceptual association support function. This finding implies that the conceptual association support helps learners to be aware of important concepts in the exploration topics since slightly more than half of the concepts in the mind maps were suggested by the concept association rules. Based on the interview, 4.77 concepts were identified as important concepts for each learner. Among these important concepts, 1.69 (SD: 1.68) were from the recommended concepts while 3.08 (SD: 1.71) were self-determined concepts. In other words, 35.4 percent of the important concepts were recommended by the conceptual association bank. This finding also suggests that the concept association recommendations can improve the learners’ understandings of the exploration domain.

However, it should be noted that the SD of the numbers of self-determined and recommended concepts in mind maps were high. This may be because the learners were in different stages of writing their theses. Among the 25 learners, 13 had been conducting the literature review for their studies while 12 had not done the literature review at the time of the experiment. The learners who were already in the literature review stage could identify more concepts to explore a domain because they already had the basic knowledge of this domain. In contrast, the learners who were not already in the literature review stage relied heavily on the concept association bank to enhance their domain knowledge. In summary, the findings from learners’ mind maps and the interview confirm that concept association recommendations are useful for expanding exploration and understanding of topics.

*Exploration activity analysis*

Analysis of the videos of learners’ mind mapping activities showed that there were three types of exploration activities: initiating direction, integrating knowledge, and seeking information (Table III).

**Initiating directions.** Learners used a variety of ways to initiate their exploration directions. Learners found a total of 418 concepts to explore a knowledge domain themselves. However, learners also frequently initiated exploration directions by the use of concept association recommendations. The learners requested concept recommendations 786 times. Among these, learners accepted concepts from the

<table>
<thead>
<tr>
<th>Note type</th>
<th>Explanation</th>
<th>Mean</th>
<th>SD</th>
</tr>
</thead>
<tbody>
<tr>
<td>Self-determined concepts in mind maps</td>
<td>The concept nodes learners specified in the final mind map</td>
<td>9.88</td>
<td>7.23</td>
</tr>
<tr>
<td>Recommended concepts in mind maps</td>
<td>The concept nodes recommended by the concept association bank in the final mind map</td>
<td>10.52</td>
<td>8.02</td>
</tr>
<tr>
<td>Total concept nodes</td>
<td>The concept nodes learners specified and the concept nodes recommended by the concept association bank in the final mind map</td>
<td>20.4</td>
<td>10.12</td>
</tr>
</tbody>
</table>

*Table II.* The distribution of different concept nodes in the mind maps of learners
recommended concept list and added the concepts into their concept maps 432 times. However, the learners deleted concepts 261 times. They also changed the exploration directions 121 times by modifying the determined concepts. This finding reflects that exploration is a dynamic process during which the learners have to iteratively seek information in some directions that they might not have known about. The concept association bank could thus help the learners explore the web by providing potential directions. Owing to suggesting new exploration directions, the concept association bank is an effective way to solve the “overspecialised” and “similarity hole” problems of the content-based recommendation systems (Adomavicius et al., 2005; McNee et al., 2006).

Integrating knowledge. Learners demonstrated rich knowledge integration activities. Learners linked concepts 715 times. Likewise, learners requested recommendations of associations between concepts 271 times. This result indicated that learners needed a facility not only to obtain information for extending exploration but also an aid to link concepts. As a result, learners largely relied on the concept association function to build the relationship between concepts. The video results indicated that learners also deleted links between concepts, which they did 306 times. The videos showed that learners deleted concepts and links in their concept maps on two occasions. The first was when learners could not find any associated concepts for a particular concept from the concept association bank. The second was when learners viewed the recommended concepts of the particular concept and found the recommended concepts were different from what they originally thought. Therefore, learners decided to delete the concepts. This finding reveals that learners may experience cognitive load (Conklin, 1989; Madrid et al., 2009) and the concept association recommendations can assist them to reduce complexity during the exploration process as it can help them to eliminate unrelated concepts in the mind maps.

Seeking information. This study found that learners frequently sought information to gain a new understanding of their exploration topics, either by browsing the web or by retrieving the list of related papers provided by the concept association bank. Web browsing behaviour is a data-driven exploration strategy to locate isolated resources on the web. However, the resources the learners found on the web were linked to the conceptual level information, i.e. concept nodes in mind maps, which were related to the exploration goal. In other words, there are two types of exploration strategies:

<table>
<thead>
<tr>
<th>Activity</th>
<th>Frequencies</th>
</tr>
</thead>
<tbody>
<tr>
<td>Initiating directions</td>
<td></td>
</tr>
<tr>
<td>Determining concepts</td>
<td>418</td>
</tr>
<tr>
<td>Requesting concept recommendations</td>
<td>786</td>
</tr>
<tr>
<td>Accepting concept recommendations</td>
<td>432</td>
</tr>
<tr>
<td>Modifying concepts</td>
<td>121</td>
</tr>
<tr>
<td>Deleting concepts</td>
<td>261</td>
</tr>
<tr>
<td>Integrating knowledge</td>
<td></td>
</tr>
<tr>
<td>Linking concepts</td>
<td>715</td>
</tr>
<tr>
<td>Requesting associations between concepts</td>
<td>271</td>
</tr>
<tr>
<td>Deleting links between concepts</td>
<td>306</td>
</tr>
<tr>
<td>Seeking information</td>
<td></td>
</tr>
<tr>
<td>Browsing the web</td>
<td>595</td>
</tr>
<tr>
<td>Retrieving a list of recommended related papers</td>
<td>870</td>
</tr>
</tbody>
</table>

Table III. Frequencies of student exploration activities
data driven and goal driven. In the data-driven strategy, learners browsed the web 595 times, starting by locating isolated resources on the web and then associating these resources with goal-level concepts. In the goal-driven strategy, the learners started by identifying a concept which was related to their exploration goal and then explored the list of related papers based on this goal. They did this 870 times. Such information seeking activities revealed that learners used the mind maps as a platform to switch between the goal-driven knowledge exploration and detailed data-driven exploration to advance their understanding.

**Sequential pattern analysis**

Figure 3 shows the transition possibilities among the main mind mapping activities during knowledge exploration on the web. The transitions with small probabilities (transition probability < 1/9, the random probability between any two of the nine mind mapping activities under uniform distribution) were eliminated so that only frequent transitions were displayed to show the learners’ main exploration behaviours. The high transition probability between the main mind mapping activities on the arcs between two activities shows the tendency of transition from one type of activity to another. One of the transitions with high probability is the transition from browsing...
the web to self-determining concepts ($p = 0.14$) which reveals that learners frequently identified a new concept for exploration after browsing the web. The transition diagram also shows that learners exhibited high transition probability from requesting concept recommendations and browsing the list of related papers to all other activities (except linking concepts). Thus, learners frequently utilised the concept association bank to help in self-determining concepts, managing mind maps and creating new understanding as detailed below.

**Self-determining concepts.** Requesting concept recommendations and retrieving the list of related papers are two major activities facilitated by the concept association bank. Thus, the analysis of the transition activities focuses on how these two activities influence other exploration activities. One of the influenced activities is self-determining concepts, which was affected by requesting concept recommendations ($p = 0.15$) and retrieving the list of related papers ($p = 0.15$). These two approaches comprised a passive approach to initiating new exploration directions. In addition to the passive approach, concepts can also be self-determined using an active approach in which the learners identify new exploration concepts through browsing the web ($p = 0.14$) and they could also identify new concepts after linking existing concepts ($p = 0.2$). In other words, learners can help themselves to obtain new exploration concepts by either seeking a potential exploration direction hidden in large information resources or managing the relationships between concepts. These findings indicate that learners cannot only take an active approach to expand exploration directions, but also seek support from the concept association bank to recommend potentially useful concept associations originally hidden in large information resources.

**Management of mind maps.** Another activity influenced by requesting concept recommendations and retrieving the list of related papers is the management of mind maps. The learners relied extensively on the concept association bank to delete concepts, delete links, and refine concepts to manage their mind maps. They frequently deleted a concept after requesting concept recommendations ($p = 0.19$) and retrieving the list of related papers ($p = 0.13$). In addition, learners sometimes decided to delete links between two concepts after retrieving the list of related papers ($p = 0.49$). Such deletions revealed that the concept association bank assisted learners to judge the usefulness of an exploration direction and its concept associations by providing related concepts and papers. Furthermore, the concept association bank helped the learners to refine exploration directions. The learners modified concepts frequently through retrieving the list of related papers ($p = 0.12$) or requesting concept recommendations ($p = 0.46$). In other words, learners repeatedly requested recommendations of related concepts and retrieved lists of related papers which in turn helped them judge the relevance of an exploration direction and its concept associations. By doing so, irrelevant concepts and associations were removed by learners based on the support from the concept association bank, helping to reduce the cognitive load and complexity during the exploration process.

**Creation of new understanding.** The learners relied on the mind maps to manage their own understanding of the exploration topics. They often added comments to their mind maps after retrieving the recommended list of related papers ($p = 0.14$). However, some learners needed additional support from the web. More specifically, the learners would take the recommended list of related papers as an initial entry point to retrieve information from the web ($p = 0.5$),
which was accumulated and integrated to generate new understanding. This new understanding was commented on ($p = 0.18$). In summary, the mind map provides different approaches for learners to advance their understanding. The learners can either directly produce new understanding or achieve this target with the support from the web resources. Thus, the mind map is a useful tool to help learners reflect on their understanding by integrating information from the web resources.

**Conclusions**

Many educators have highlighted the importance of using online resources to extend individual learning. Support to facilitate knowledge exploration on the web is important when learners only have limited knowledge of a subject domain. This study proposed a conceptual association bank that utilises a mind map as a knowledge integration platform and extracts concept associations from collective domain databases to support learners to explore a subject domain. Figure 4 shows the conceptual framework of expanding understanding based on the findings of this study.

As shown in this framework, the concept association bank and mind map approaches can help the learners to expand their understanding of a subject domain and this goal can be achieved with a variety of sub-goals:

- *Providing exploration direction*. This study found two approaches to finding directions in which to explore a subject domain. Some learners found a new direction by browsing the web independently while most of the learners relied on the community intelligence provided by the concept association bank in this study.
• **Reducing complexity and cognitive load.** The learners’ exploration activities suggested that there is a need to help the learners manage complexity and reduce their cognitive load. The concept association bank is useful in helping learners to fulfil this need through deleting/modifying concepts and deleting links between concepts.

• **Facilitating data- and goal-driven exploration strategies.** Learners’ exploration activities demonstrated that they not only seek information on the web (data-driven strategy) to form a new concept but also take an unfamiliar concept and then elucidate its meaning by seeking information on the web (goal-driven strategy). It was found that the concept association bank facilitated learners to apply the goal-driven strategy by providing potential related concepts for further exploration. In addition, the mind maps facilitated learners to integrate concepts from individual resources and switch between the goal- and data-driven strategies of exploration.

• **Commenting on new understanding.** The learners’ exploration activities showed two pathways to commenting on new understanding. While some learners directly commented on a resource on the mind map, most of the learners, starting from the recommended list of related papers, iteratively browsed the web to achieve and comment on new understanding of the subject domain.

Validity is an issue which needs to be addressed in further work. The experiments were conducted using the mind map system proposed by this study, but the four main conclusions detailed above should be valid for other settings using other knowledge integration systems such as FreeMind and Inspiration because of the standard sets of functionalities that the mind map system offers to learners. For the same reasons the design guidelines given for the systems which facilitate knowledge exploration should apply to any knowledge integration systems. We also think that, to a large extent, the main conclusions are likely to be valid for general information seeking settings such as searching with Google or Yahoo. This is because seeking information on the internet is similar to the knowledge exploration behaviour of the participants in this study. However, further studies are required to verify the findings of this study and adjust the framework of expanding understanding in such settings.

The results of this study demonstrate that a mind map combined with the association bank is a useful resource for student learning. However, this study was only a small-scale investigation. Further work needs to be undertaken with a larger sample to provide additional evidence. In addition, this study only used TF-IDF and association rule techniques to generate the concept association bank. It would be interesting to see what results would be generated by using other methods such as a support vector machine or Bayesian techniques. Gathering information on these issues through further work can help clarify the findings from this study. In addition, the results of such studies could be integrated to build sophisticated facilities to help learners explore knowledge in large databanks.

**References**


About the authors
Chen-Chung Liu is an Associate Professor in the Graduate Institute of Network Learning Technology, National Central University, Taiwan. He leads the iLearn group (Interactive Technology for Learning) which aims to develop socio-technical environments that augment individual learning and creativity in classrooms and informal settings. Chen-Chung Liu is the corresponding author and can be contacted at: ccliu@cl.ncu.edu.tw

Shih-Hsun Fan Chiang was a graduate student in the iLearn group.

Chih-Yueh Chou is an Assistant Professor at Yuan-Ze University.

Sherry Y. Chen is a Professor in the Graduate Institute of Network Learning Technology, National Central University, and a Reader in the Department of Information Systems and Computing at Brunel University, London. She obtained her PhD from the Department of Information Studies, University of Sheffield, UK. Her current research interests educational technology, human-computer interaction and data mining.

To purchase reprints of this article please e-mail: reprints@emeraldinsight.com
Or visit our web site for further details: www.emeraldinsight.com/reprints