Assessing Conceptual Similarity to Support Concept Mapping*

David B. Leake and Ana Maguitman
Computer Science Department
Lindley Hall, Indiana University
150 S. Woodlawn Avenue
Bloomington, IN 47405, U.S.A.
{leake,anmaguit}@cs.indiana.edu

Alberto Cañas
Institute for Human and Machine Cognition
University of West Florida
11000 University Parkway
Pensacola, FL 32514, U.S.A.
acanas@ai.uwf.edu

Abstract

Concept maps capture knowledge about the concepts and concept relationships in a domain, using a two-dimensional visually-based representation. Computer tools for concept mapping empower experts to directly construct, navigate, share, and criticize rich knowledge models. This paper describes ongoing research on augmenting concept mapping tools with systems to support the user by proactively suggesting relevant concepts and associated resources (e.g., images, video, and text pages) during concept map creation. Providing such support requires efficient and effective algorithms for judging concept similarity and the relevance of prior concepts to new concept maps. We discuss key issues for such algorithms and present four new approaches developed for assessing conceptual similarity for concepts in concept maps. Two use precomputed summaries of structural and correlational information to determine the relevance of stored concepts to selected concepts in a new concept map, and two use information about the context in which the selected concept appears. We close by discussing their tradeoffs and their relationships to research in areas such as information retrieval and analogical reasoning.

Introduction

Capturing expert knowledge is an essential component of the knowledge management process. Once models of experts' domain knowledge are available, they can provide a valuable resource for knowledge comparison, refinement, and reuse. However, a difficult question is how to obtain the required knowledge models. Hand-crafting is expensive; machine learning techniques may not be effective. We are investigating an alternative approach: developing tools to enable experts themselves to construct models of their knowledge. Our approach builds on *concept mapping* (Novak & Gowin 1984), in which subjects construct a two-dimensional, visually-based representation of concepts and their relationships. Concept mapping was first proposed in educational settings, to help assess students' understanding and to aid their knowledge-building, comparison, and refinement. In the concept mapping view, experts who build concept maps are not simply externalizing pre-existing internal knowledge, but are also doing knowledge construction. Thus tools to provide relevant knowledge to consider and compare during concept mapping could facilitate not only knowledge capture, but knowledge generation.

The Institute for Human and Machine Cognition has developed a set of publicly-available tools for concept mapping, available at http://cmap.coginst.uwf.edu/. These widely-used systems support generating and modifying concept maps in electronic form, as well as annotating concept maps with additional material such as images, diagrams, and video clips. They provide the capability to store and access concept maps on multiple servers, to support knowledge sharing across geographically-distant sites. We have developed an initial implementation of a suggester system that automatically extracts information from a concept map under construction and uses that information to retrieve prior concept maps, associated resources, and related concepts that the user can compare and possibly include in the concept map being constructed. Figure 1 shows a screen shot of the concept mapping tools being used for knowledge modeling about Mars, with the suggester proposing new concepts to link to the "space exploration" node (to fill in the not-yet-specified concept node designated by "????").

The effectiveness of a suggester system depends on efficient algorithms for judging similarity and relevance of stored concepts to the concepts currently under consideration. This paper describes four approaches that we have implemented and are now testing, two of which focus on determining the relevance of a prior concept to a new concept, based on summaries of structural and correlational information previously generated for the concept map library, and two of which directly compare the context in which the concept appears—its concept map—to concept maps in the concept map library. We compare the complexity of these approaches, discuss pilot studies of their effectiveness, and the relationship of this work to previous approaches.

Concept Maps for Knowledge Modeling

Concept mapping was designed both to enable the examination of human conceptualizations, and to further human knowledge construction. As shown in the center of Figure 1, concept maps are a two-dimensional visual representations

* This research is supported in part by NASA under award No NCC 2-1216.

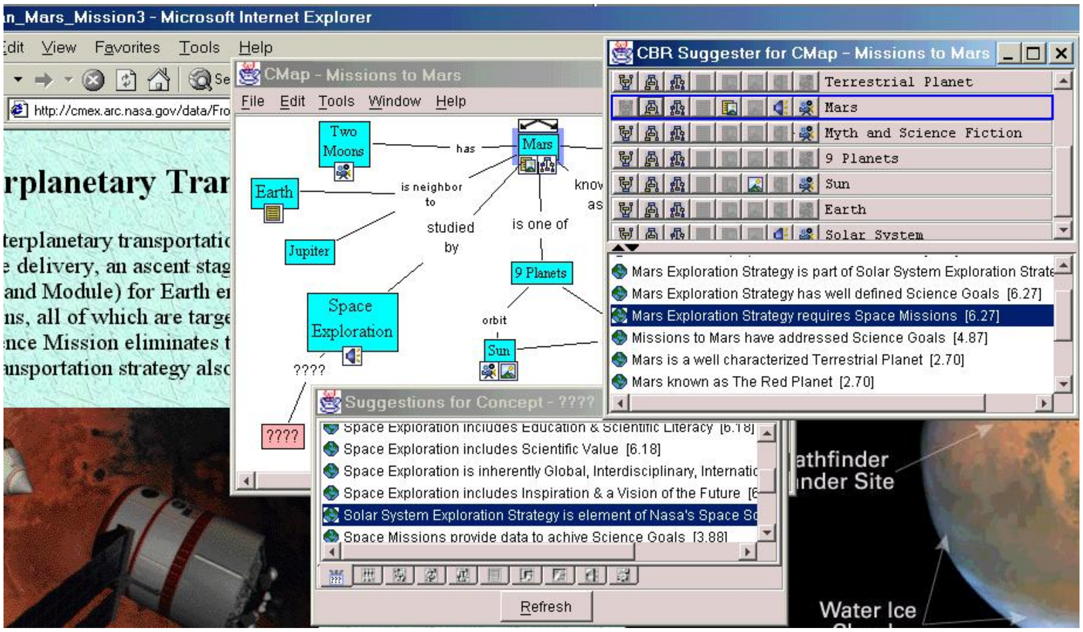

Figure 1: A screen shot of the suggester proposing resources relevant to a current concept, in the context of a concept map.

containing nodes for concepts, connected with named links expressing concept relationships (e.g., that Earth is a neighbor of Jupiter). Concept maps appear similar to semantic nets but have no fixed semantics and vocabulary—they simply make explicit any set of concepts and relationships in any vocabulary that the expert chooses.

In electronic concept maps, nodes can be associated with resources such as photographs and textual passages (as shown in the background of Figure 1), diagrams, or even pointers to additional concept maps to define a hierarchical concept structure. The result is a rich and flexible concept representation to help humans understand domains and revise their domain knowledge. Concept mapping has received widespread use for knowledge modeling, sharing, and refinement by experts and novices (e.g., in the Quorum project, involving over a thousand schools in South America (Cañas *et al.* 1995)). As increasing numbers of concept maps are captured in electronic form, they provide a growing source of data for studying human concepts, for enabling knowledge sharing, and for helping to refine tools to support human concept-mapping.

Some Central Issues

Developing methods for assessing concept map similarity requires addressing issues for both cognitive science and AI:

- **The roles of content and structure in similarity assessment:** Models of conceptual similarity in concept maps must consider both concept labels, and how the labeled concepts are related to other concepts.
- **Assessing similarity and relevance for non-standardized representations:** Labels on concept map nodes provide names for the concepts that they represent, but not in the more formal, standardized representations assumed in much AI research. Node and link labels may be ambiguous or inconsistent with the names used in other concept maps. Thus determining related concepts requires more than simple keyword matching, and similarity assessment must be sufficiently robust to deal with representational differences.
- **Efficient use of structural information**: If link labels cannot be matched reliably, matching concept map structure reduces to graph matching. Because this is expensive, methods are needed to summarize structural information and use those summaries to guide matching.
- **Exploiting contextual information:** Context may be crucial in determining the relevance of concepts with different labels, because the meaning of each concept in a concept map is partially captured by its connections to other concepts. Context may be crucial even for determining relevance of identical concepts: A rocket engine designer who enters a node labeled "hydrogen peroxide," linked to concepts for fuel and propulsion, would not be interested in retrieving a concept map on first-aid that happens to include hydrogen peroxide as well.
- **Facilitating representational standardization:** The usefulness of conceptual information for reasoning systems increases with standardization. To increase standardization without increasing the burden on users, methods are needed to help identify to reuse existing labels.

Methods for Computing Relevance, Similarity, and Usefulness

We are studying techniques for assessing the relevance of a new concept as a candidate "concept extension" (related concept to consider linking to a selected concept in a concept map), as well as to suggest relevant vocabulary items for possible reuse. This section introduces and compares four techniques for determining the relevance of a new concept to a concept under consideration. The first two rely primarily on precomputed global information, while the second two use the context of the concept map in which the concept appears. We begin with definitions that will be useful for understanding the following algorithms.

Preliminary Definitions

Concepts have different importances in concept maps, and the concept map layout often provides useful information for assigning concept weights. For example, a main concept usually appears at the top of each concept map, specifying the main topic. In (Cañas, Leake, & Maguitman 2001) we proposed that a small set of topological dimensions can usefully summarize concept roles:

- *Authorities:* Concepts to which other concepts converge. These have the largest number of incoming links from "hub nodes" (defined below).
- *Hubs* (centers of activity): Concepts with the largest number of outgoing links ending at *authority nodes*.
- *Upper Nodes:* Concepts that appear towards the top of the map in its graphical representation.
- *Lower Nodes:* Concepts that appear towards the bottom of the concept map in its graphical representation.

Our algorithms to compute these weights are adapted from research on determining hub and authorities nodes in a hyperlinked environment (Kleinberg 1999). We define four weights, *a-weight*, *h-weight*, *u-weight* and *l-weight*, in [0,1], representing the degree to which a concept belongs to the above categories in a particular concept map. Detailed definitions are presented in (Cañas, Leake, & Maguitman 2001). For a given concept map, these weights can be computed in $O(n^3)$ time, where n is the number of concepts in a map. They need only be computed once, when the concept map is indexed, and stored with each concept.

To describe individual concepts, our methods extract keywords from the concept labels ("stop words" are deleted before processing), and weight the keywords in terms of these four types of weights, using the weights of the concepts in which they appear. Concept maps are then compared according to their weighted keywords. (In the following formulas, we sometimes refer to applying set operations such as intersection and difference to concept maps; this is a shorthand for applying those operations to the sets of keywords extracted from the concept maps.) Given a keyword k and concept map library L, Θ_k^L stands for the set of concept maps in L containing keyword k. For simplicity we assume that L is fixed and use Θ_k to denote all concept maps containing keyword k. It may be useful to refer to the global weight of a keyword in a set of concept maps. If Θ is a set of concept maps, k is a keyword, and *w* is a weight function, the global weight $\Upsilon^w(\Theta, k)$ of k in Θ is defined by:

$$\Upsilon^w(\Theta, k) = \sum_{C \in \Theta} w(k, C).$$

Some of our algorithms compute the average of the n highest values of a set of values. For a set of values A, we will use the notation $\Xi(n, A)$ to refer to the sum restricted to the n highest values of A, divided by n. In the special case when A is empty, the returned value is 0.

Estimating Relevance by Global Correlations

Our first two methods use global correlation metrics to retrieve concept maps containing concepts that tend to co-occur with concepts from the current concept map. Because this allows correlated keywords to match each other, it is more flexible than using keyword matching alone. Correlation information is combined with the weight each keyword has on the corresponding maps—giving rise to *weight-based* global correlation metrics—or with the distance between the two involved keywords in each concept map—giving rise to *distance-based* global correlation metrics. (In a concept map the notion of distance can be naturally defined as the minimum distance between concepts in which the keywords appear.) Both of the above methods are global in the sense that they assign concept importances based on global information pre-computed from the entire concept map library.

Method 1: Using weight-based global correlations: To compute *weight-based global correlations* between a source concept S and a target concept T, we first compute the set $WV(S,T)$ of *weight-based correlation values*. Writing Υ for $\Upsilon^{u-weight}$, we calculate:

$$\left\{ \frac{|\Theta_i \cap \Theta_j|^2 * \Upsilon(\Theta_i \cap \Theta_j, i) * \Upsilon(\Theta_i \cap \Theta_j, j)}{|\Theta_i| * |\Theta_j| * \Upsilon(\Theta_i, i) * \Upsilon(\Theta_j, j)} : i \in S, j \in T \right\}.$$

Then we compute the *weight-based global correlation* as $\mathcal{W}(S,T) = \Xi(n, WV(S,T))$, where $n = (|S| + |T|)/2$.

Method 2: Using distance-based global correlations: In order to compute *distance-based global correlations* we start by defining $D_C(i,j)$, which states the distance between keywords i and j in the concept map C. The distance metric $D_C(i,j)$ can be naturally defined as the minimum number of links between concepts containing those keywords, or infinity if i and j are not both in C. Consider the set of keywords S and T. We begin by computing the set $MV(S,T)$ of *distance-based correlation values*:

$$\left\{ \sum_{C \in (\Theta_i \cap \Theta_j)} \frac{2}{(|\Theta_i| + |\Theta_j|) * D_C(i,j)} : i \in S, j \in T \right\}.$$

The *distance-based global correlation* is then $\mathcal{M}(S,T) = \Xi(n, MV(S,T))$, where $n = (|S| + |T|)/2$.

Estimating Relevance Based on Context

When providing suggestions to a user constructing a concept map, it is appealing to retrieve concepts that appear in contexts similar to the map under construction. We have developed two methods to compare local contexts, the first of which is a *similarity-based* approach. It posits that the more similar the contexts in which two concepts appear, the higher the likely relevance of one concept to the other. The second is a *usefulness-based* approach that favors concepts providing additional information as the user seeks to add new concepts to a partial concept map.

Method 3: Using contextual similarity: To compare the concept map structures in which two concepts appear, we consider the four weights obtained from topological analysis to summarize the positioning of each concept, and use that information to compare the role of each related concept (i.e., concept with overlapping keywords) in its own concept map by calculating the distances between the sets of associated weights. Thus the weights obtained from the topological analysis of a concept map are used to define *topological similarities* between two concepts belonging to two different concept maps. For example, two concepts that appear at the top of their corresponding concept maps will have similar *u-weights*, while two concepts that play similar roles as *hub nodes* in their corresponding concept maps will have similar *h-weights*. Based on these intuitions we can compare two maps C_S and C_T by first computing $TS(C_S, C_T)$, the set of *topological similarity values* as follows:

$$\left\{ \frac{|S \cap T| * \sum_{\Delta w \in \Delta W} (1 - \Delta w(S, C_S, T, C_T))}{|\Delta W| * \alpha(S, T)} : S \in C_S, T \in C_T \right\}$$

where

- $\alpha(S, T) = (|S| + |T|)/2$,
- $\Delta W = \{\Delta$*a-weight*$, \Delta$*h-weight*$, \Delta$*u-weight*$, \Delta$*l-weight*$\}$, and
- $\Delta w(S, C_S, T, C_T) = (w(S, C_S) - w(T, C_T))^2$, and w is one of the four weights.

We then compute the *topological similarity* between concept maps C_S and C_T, as $\mathcal{S}(C_S, C_T) = \Xi(n, TS(C_S, C_T))$, where $n = (|C_S| + |C_T|)/2$.

Method 4: Using context and novelty of information: The concept map that is the most similar to the source map may not be most useful for suggesting information to connect to a new concept. For finding new connections, useful prior concept maps are those that both include similar concepts and suggest new connections. Consequently, we are also exploring methods that favor both commonality and the existence of new material in the stored concept map.

A simple usefulness measure between a source concept map C_S and a target concept map C_T can be computed by:

$$\mathcal{U}(C_S, C_T) = \alpha * |C_S \cap C_T| + \beta * |C_T - C_S| - \gamma * |C_S - C_T|$$

where α, β and γ are constants that adjust the balance between overlap and novelty. ($\mathcal{U}$ need not be symmetric, so it is a *measure*, rather than a metric.) We are also investigating measures that consider the correlations between target and the source keywords for more flexible matching of non-identical terms. For example, applying the distance-based correlation measure from Method 2, we can compute the usefulness measure:

$$\begin{aligned}\mathcal{U}_{\mathcal{M}}(C_S, C_T) = \ & \alpha * |C_S \cap C_T| + \\ & \beta * \mathcal{M}(C_S, C_T - C_S) * |C_T - C_S| - \\ & \gamma * (1 - \mathcal{M}(C_T, C_S - C_T)) * |C_S - C_T|.\end{aligned}$$

Discussion of Methods

Assessing the previous methods requires considering their cost and the quality of their relevance predictions.

Cost: Methods 1 and 2, the global correlation metrics, are efficient to compute. Computing the global weight for a keyword (Method 1) is linear in the number of concept maps involved. Computing the weight-based correlation values for keywords i and j involves counting the concept maps simultaneously containing those keywords, which can be done in times from $O(|\Theta_i|)$ to $O(|\Theta_i| * |\Theta_j|)$, depending on the underlying indexing scheme. Weight-based global correlations must be computed for each pair of keywords in source and target concepts, but because the number of keywords in concepts is usually small, this is inexpensive in practice. Computing distance-based global correlations (Method 2) requires computing minimum distances between pairs of keywords in concept maps, which is basically the shortest path problem, and can be computed in $O(|E| * log|V|)$, where V is the number of vertices (concepts in our case) and E is the number of edges (links in our case).

Method 3, the first context-based method, requires computing topological similarity between each pair of concepts in two concept maps. In principle, this can be quite expensive, but these values only need to be computed for pairs of concepts that have at least a keyword in common. Usually there are few such concepts in any concept map, making this inexpensive in practice. Depending on the indexing mechanism used, the basic technique considering context and novelty (the first version of Method 4) can be implemented in times $O(m)$ to $O(m * n)$, where m and n are the sizes of the concept maps to be compared. The second version of Method 4, which adds global correlations, is significantly less efficient. Its speed of calculating global correlations is reasonable for comparing individual concepts, but not for comparing concept maps. In future research we intend to perform a formal analysis and to develop efficient approximations of this approach.

Relevance of suggestions: We performed a pilot experiment to evaluate whether our metrics can be exploited to provide better recommendations than the simple baseline method of counting shared keywords. In the experiment, subjects were presented with a concept map, with one concept designated as the concept to be extended, and a list of 50 suggestions chosen randomly from the set of extensions containing at least one keyword in common with the concept to be extended. Ten subjects, all graduate students not involved in the project, assessed the relevance of retrieved information on a scale of 0 to 10. Their rankings were compared to the relevance scores assigned by our techniques. We then used Spearman rank correlation to compare the ranking produced by the human subjects to the ranking produced by our algorithms.

In our study, best results are obtained when usefulness-based comparison measures (Method 4) are used to pre-select target concept maps and the combination of weight-based comparison metric (Method 1) and distance-based comparison metric (Method 2) is used to rank suggestions. In this approach, the weight-based metric measures the similarity between the source base concept, i.e., the concept to be extended, and the target base concept, i.e., the concept that is connected to potentially relevant extensions; the distance-based metric compares the base source concept to the potentially relevant extension. Intuitively, weight-based metrics between concepts tell us how similar two given concepts are, while the distance-based metric helps to predict how suitable is for a concept to have another new concept as a neighbor.

Our results show a correlation factor of 0.77, with a 2-tailed significance level $p < 0.0001$, between the values obtained by the combined method and the aggregation of the evaluations made by human subjects. Our results show a correlation factor of 0.63, with $p < 0.0001$, between values from the baseline "counting common keywords" and the aggregate of the results returned by human subjects. This suggests that our methods are capturing regularities beyond those captured by the baseline method.

Comparison to Related Work

The project described here relates to numerous research areas such as knowledge modeling and sharing, concept representation, information retrieval, and case-based reasoning. As a knowledge modeling project, it contrasts with knowledge engineering approaches that depend on hand-crafting knowledge representations, aiming to empower domain experts to directly construct, navigate, share, and criticize knowledge models. This requires that the concept representations be natural for them to construct, and sufficiently expressive for others to understand their conceptualizations. The representation that we have chosen, concept maps augmented with supplementary resources, appears to provide the needed information in an easy-to-use form, and a recent study substantiates the usefulness of concept map navigation for guiding knowledge access (Carnot *et al.* 2001).

Keyword-based retrieval techniques are common in the information retrieval literature (Baeza-Yates & Ribeiro-Neto 1999). Because concept maps provide additional structure, our methods augment keyword-based methods with consideration of the topological role of a keyword, inherited from the topological role of the concept in which it appears. The IR community has also done considerable research on methods involving *metric clusters*, in which keywords are compared in terms of their distance (usually defined by the number of words between them in a document, with infinite distance between keywords in different documents). Our notion of *distance-based global correlations* is an adaptation of these ideas.

The comparison of structured information has been extensively studied in research on case-based reasoning and analogical reasoning. Our methods rely on topological analysis techniques rather than on explicit structure mapping, as done by (Falkenhainer, Forbus, & Gentner 1989). Structural analysis requires a standardized representation language, and assumes that the most important matches involve the links, rather than the entities that these links relate. In concept maps, the representational vocabulary is nonstandardized and link names tend to be generic, so the most significant information source is usually on the concepts rather than on the links.

Conclusion

Concept mapping provides a means to capture and examine human concepts, as well as a tool for aiding experts and novices at constructing and refining their own understanding of a domain. Augmenting concept mapping tools with intelligent methods for suggesting relevant concepts to compare and consider is promising for aiding these processes and facilitating knowledge sharing. Developing these methods depends on being able to efficiently and effectively assess the relevance of concepts in prior maps to selected concepts in the concept maps currently being constructed.

The paper has identified key issues for this task and presented a set of approaches for assessing conceptual similarity and relevance for concept mapping. These approaches have been implemented in a suggester system combined with the electronic concept mapping tools of the Institute for Human and Machine Cognition, with encouraging initial results that we are preparing to test more extensively in a larger-scale study. Based on the results of that study, we intend to refine these individual methods and investigate possibilities for combinations to exploit their individual strengths, as well as to address additional issues such as selectively adjusting concept weights to reflect additional information about task contexts.

References

Baeza-Yates, R., and Ribeiro-Neto, B. 1999. *Modern Information Retrieval*. Addison-Wesley.

Cañas, A.; Ford, K.; Brennan, J.; Reichherzer, T.; and Hayes, P. 1995. Knowledge construction and sharing in quorum. In *World Conference on Artificial Intelligence in Education, AIED'95*, 218–225. AACE.

Cañas, A.; Leake, D.; and Maguitman, A. 2001. Combining concept mapping with cbr: Experience- based support for knowledge modeling. In *Proceedings of the Fourteenth International Florida Artificial Intelligence Research Society Conference*, 286–290. AAAI Press.

Carnot, M.; Dunn, B.; Cañas, A.; Graham, P.; and Muldoon, J. 2001. Concept maps vs. web pages for information searching and browsing. http://www.coginst.uwf.edu/users/acanas/Publications/-CMapsVSWebPagesExp1/CMapsVSWebPagesExp1.htm.

Falkenhainer, B.; Forbus, K.; and Gentner, D. 1989. The structure-mapping engine: Algorithm and examples. *Artificial Intelligence* 41:1–63.

Kleinberg, J. 1999. Authoritative sources in a hyperlinked environment. *Journal of the ACM* 46(5):604–632.

Novak, J., and Gowin, D. 1984. *Learning How to Learn*. New York: Cambridge University Press.