CASE BASED CONCEPT MAP TOPOLOGY COUNSELOR

Sofía Brenes & Alejandro Valerio, Computer Science Department, Indiana University.
Email: {sbrenesb, avalerio}@cs.indiana.edu

Abstract. In the process of constructing a concept map, a common problem is how to achieve a correct topological distribution of the concepts created. The project described in this paper uses Case Based Reasoning techniques to develop a Concept Map Topology Counselor that suggests improvements to a concept map based on its layout. The case base is constructed using examples of concept maps that present topological problems such as a tree structure (no cross-links), concepts with too many outgoing links, "String", and unbalanced concept maps. Each case has a corresponding solution in the case base, consisting of the same concept map with concepts and linking phrases either added or removed. A modified graph isomorphism algorithm is used to compare a concept map with the ones present in the case base. If a match is found, the solutions are shown to the user who can then decide whether or not to incorporate the changes. In this paper we describe the implementation of a prototype as well as the results of preliminary evaluation of the Counselor.

1 Introduction

In the process of constructing a concept map (Novak & Gowin, 1984), people sometimes struggle to find an appropriate distribution of concepts and linking phrases, leading to concept maps with poor topological layouts. Automatic systems have been successful in making recommendations to final users by analyzing the content or structure of concept maps (Leake et al., 2004; Cañas et al., 2004a).

Suggesting improvements to a concept map's topology has been previously addressed in CmapTools (Cañas et al, 2004b) from the Institute for Human and Machine Cognition through the “Joe in a Box” tool, which applies a set of heuristics to find layout problems in a concept map. These heuristics were collected by observing common layout “mistakes” made by starter users when constructing a map. Its drawback is that these hard coded rules are applied indiscriminately, so even if a rule is found it does not always mean that the map has a problem. There are also rules that fail to capture other problems, limiting the power of the tool and rendering it inaccurate sometimes.

In this paper we present the prototype of an automatic system that uses Case Based Reasoning (CBR) (Kolodner & Leake, 1996) to identify potential topological problems in concept maps and suggest alternative distributions based on the similarity of the current concept map to the existing set of cases, rather than fixed rules. By applying a more flexible approach, we intend to overcome some of the limitations presented by “Joe in a Box”. Unlike other recommendation systems that focus on the content of concepts and linking phrases, this work centers on the general layout of the map, as an alternative to provide automatic assistance to users in the construction process.

The paper is structured as follows: Section 2 details the proposed solution and initial implementation of the system, covering the case base construction and retrieval algorithm. Section 3 contains a preliminary evaluation and its results while Section 4 provides a summary and intended future work.

2 Concept Map Topology Counselor

CBR is a machine learning process that relies on analogy to find solutions in a collection of past experiences or cases. When a new problem is presented, it is matched against the collection, the most similar cases are retrieved, and their solutions are used, and sometimes adapted, to give an answer to the new situation. The original problem combined with the given solution become a new case and could be integrated to enrich the case base collection.

In our proposal, a case is a concept map with some topological problem and the corrected version of the map is the solution (equivalent to the problem-space and solution-space in a CBR system respectively). A coloring scheme is used in both maps to tag the nodes exhibiting the problem and their corresponding solution. This tagging is also used as an index to search the problem-space, as is described next.

To find topological anomalies, a modified subgraph isomorphism algorithm is used to compare the user's concept map to the colored sections in each of the concept maps from the problem-space. A match will indicate that a problem does exist, in which case the corresponding solution, i.e. the corrected version of the concept map, is presented. When browsing each of the possible solutions, the user's concept map is highlighted in the sections that matched the concept map in the problem-space. The system does not adapt the concept map automatically, instead the user decides whether to make the changes using the presented solution as a guide.

With our current proposal, the topological problems that can be identified by the Counselor are a subset of those originally modeled in "Joe in a Box". These are:

- *Tree Structure*: A tree structure is found when there are no "cross-links" between any concepts, in which case the user may not have realized there is related information in their concept map.
- *Outgoing Links*: When a linking phrase has too many outgoing links there might exist another relation that should be placed between the concept in question and the concepts that the links point to.
- *"String"*: String concept maps occur when the user has linked together concepts in a line, without any other links other than the ones that join them. This could mean the user is forming a sentence within a concept map.
- *Unbalanced*: Occurs when one side of the concept map has more depth than another. Then, the information is being focused only on a particular area of the topic.

2.1 Case Base Construction

To construct an initial case base, eighty concept maps showing some topological anomaly were collected from the IHMC Public Cmaps Server. It was desirable to have maps created by end users, rather than an artificial collection that could be biased to what the system was trying to accomplish. Each of the maps was manually corrected and both versions were marked in a color scheme that identified problems and solutions as previously described. The case base was formed by the problem-space which contains the concept maps that exhibited the anomalies, and the solution-space that contains the proposed solution to each of the problem cases.

The coloring of the concept maps served two purposes: (1) traceability of the changes from the original to the modified version and (2) indexing of the problem-space. The coloring is used by a subgraph isomorphism algorithm to detect similarities between the user's concept map and the ones in the problem-space during the retrieval process. This allows the cases to be generic, without specific or hard coded rules. All the indexed cases used in the implementation can be found in the "B552 Case Base" folder in the IHMC Public Cmaps Server.

2.2 Retrieval Algorithm

Following the usual definition of subgraph isomorphism, in Figure 1 graph A is an isomorphism of graph B. However, this does not work for our problem, since the graphs A and B would represent concept maps with very different topology.

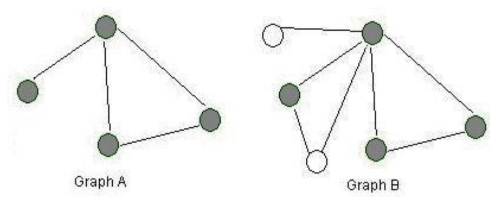

Figure 1: Subgraph Isomorphism Example.

For this reason, the definition of subgraph isomorphism used for this problem is the following: A concept map A is considered a subgraph of a concept map B if (1) A is fully contained in B and (2) each of the nodes in A has the same number of incoming and outgoing links as its corresponding node in B.

When comparing the user's concept map with the case base, we consider only the colored section of each map in the problem-space, as this is where the anomaly was identified. The colored section of a map is called the "index graph" and the user's concept map is called "target". The following algorithm is executed for all index graphs. First, a spanning tree S of the index graph is produced to ensure that each node is visited only once. Next, a combinatorial process assigns candidate nodes from the target into nodes of S. A node is a valid candidate if it has the same number of incoming and outgoing links as the node in S and if it is connected to a previously assigned candidate node. These two conditions significantly reduce the search space. The iteration process stops when all nodes in S have a candidate node assigned. After this, a full check of the edges is done to determine if the current assignment is a valid mapping. The process continues until a successful mapping is found or all combinations are exhausted.

3 Evaluation and Results

The architecture of the IHMC CmapTools software was used to add our Concept Map Topology Counselor as a new module in the application. Existing concept map data structures were used to handle the case base and obtain concept map data. Figure 2 shows the Topology Counselor as it would appear to the user.

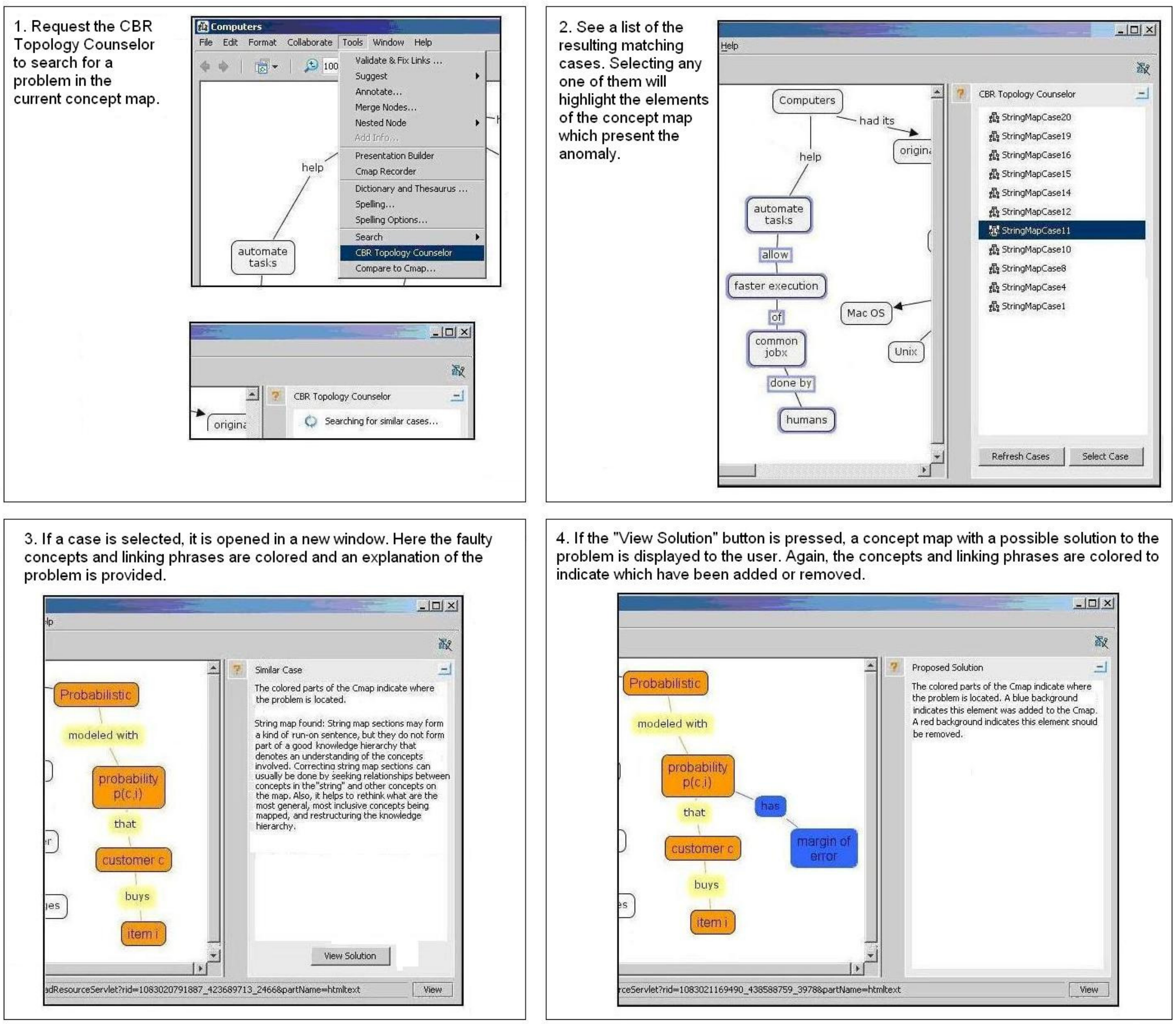

Figure 2: The Concept Map Topology Counselor.

Preliminary tests were conducted by presenting the system with concepts maps in the IHMC Public Cmaps Server that were not in the case base. This included concept maps that showed topology problems and some that did not.

We found that the system successfully identified topological problems by matching cases in the case base and correctly recognizing the entities (concepts or linking phrases) showing the problem. Even when more than one problem was present, it found all matching cases.

In the majority of tests, the efficiency of the algorithm surpassed our initial expectations given its potential combinatorial complexity. However, as the size of a concept map increases, the efficiency of the algorithm degrades rapidly failing to provide results in a reasonable amount of time. This does not happen with an average sized map, but only in very large ones, making use of the Counselor feasible for most situations.

4 Summary and Future Work

This paper presents a different approach to detect and solve topological problems in concept maps. This approach does not involve explicit rule definition. It allows a greater degree of flexibility since the case base can be expanded to widen the range of problems that can be addressed without the need of changing the design or implementation of the system. Additionally, by adding more instances of the same kind of anomaly the Counselor can more effectively recognize concept maps that present it. Unlike other approaches that use the information contained in concepts and linking phrases, this work focuses on the general topological layout of the map, as an alternative to provide recommendations to the user. Initial testing produced encouraging results and leads us to believe it is an approach worth of further study.

Future work may include the implementation of the adaptation phase, in which the concept map would be automatically corrected using the solution-space cases and giving the user the option to reject the changes suggested. Also, the retrieval algorithm must be revised in order to improve its execution response time. Although we use heuristics to reduce the search space, there are still cases in which the number of possible combinations becomes large, considerably slowing the Counselor's response time. Most importantly, the system needs to be evaluated using a formal experiment to measure its real performance.

It is also necessary to address the problem of case base maintenance. At this stage, the system does not provide a procedure to add or remove cases. This is an important issue because by adding more types of problems the accuracy of the system would increase, providing more useful recommendations to the user. The addition of new cases should be supervised by an expert to ensure their quality, instead of allowing the final user or the system itself to add the cases automatically.

5 Acknowledgements

We thank Dr. David Leake for his input throughout the development of the prototype and also Dr. Joseph Novak, Dr. Alberto Cañas, and the IHMC CmapTools team.

6 References

Cañas, A. J., Carvalho, M., Arguedas, M., Leake, D. B., Maguitman, A., Reichherzer, T, (2004a). Mining The Web to Suggest Concepts During Concept Map Construction. Proceedings of the 1st International Conference on Concept Mapping. Pamplona, Spain.

Cañas, A. et al. (2004b). CmapTools: A Knowledge Modeling and Sharing Environment. Proceedings of the 1st International Conference on Concept Mapping. Pamplona, Spain: Universidad Pública de Navarra.

Leake, D., Maguitman, A., Reichherzer, T. Cañas, A., Carvalho, M., Arguedas, M., Eskridge, T. (2004). "Googling" from a Concept Map: Towards Automatic Concept-Map-based Query Formation. Proceedings of the First International Conference on Concept Mapping, Pamplona, Spain.

Kolodner, J. L. & Leake, D. (1996). A Tutorial Introduction to Case-Based Reasoning . In Case-Based Reasoning: Experiences, Lessons and Future Directions. MIT Press. pp 31-65

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