Project F: Proceeding Paper

Comparison of Process Maps in a Diagnostic Skills Training Project

Haisong Ye

Indiana State University

CIMT 640

April 19th, 2013

Comparison of Process Maps in the Diagnostic Skills Training Program

Haisong Ye
Department of Curriculum, Instruction and Media Technology
Indiana State University
United States
hye@sycamores.indstate.edu

Abstract: This paper will focus on a method that is used to compare a process map developed by students with an expert's map and return instant feedback to students in the diagnostic skills training project. A graph comparison algorithm, similarity-flooding algorithm, and a semantic-based text similarity algorithm are used to compare the process maps. It is a challenge to gain a high accuracy score or degree on how much two maps are similar, because only humans can understand the exact meaning of a map in a particular situation. A field test with a small group of students was conducted to check the validity and reliability of the method. Discussions and suggestions are made for improvement.

Introduction

In the "Advancing Diagnostic Skills Training in the Undergraduate Technology and Engineering Curriculum" project, researchers want to develop diagnostic skills within learners who have limited content expertise. The experimental process used in this project will expose students to process mapping and then allow them to develop process maps to address simulated technical problems. After the students map out their proposed process, their map will be compared with an expert's process map. It is hoped that the mapping process and expose to expert reasoning will result in improved diagnostic thinking in general.

Process map is a kind of visual map. Visual map, which is also called mind map, has been defined as "visual, non-linear representations of ideas and their relationships" (Davies, 2011, p. 281). Mind map is more general than concept map. Typically, concept map consists of a set of propositions, which are made of nodes and connections, has "a hierarchical 'tree' structure with super-ordinate and subordinate parts" (Davies, 2011, p. 283). Process map that is used in the training program is a mind map to describe a plan of process, especially a plan of diagnostic/trouble shooting process. The particular characteristics of a process map in this paper are: (a) it has directed links, (b) it has open-ended nodes, and (c) the order of nodes is important.

The focus of this paper will be on the problem of how well a computer program can compare a learner's process map with an expert's process map and give the learner positive feedback using a process known as similarity flooding.

Literature Review

In the past decade, researchers were interested in using concept map as a strategy for assessment (Williams, 2004) in variety of subjects, such as psychology courses (Jacobs-Lawson & Hershey, 2002), computer programming (Keppens & Hay, 2008), problem-based medical curriculum (Kassab & Hussain, 2010), and so on. Results from these researchers show that concept maps are a valid means of assessment. However, inefficient manual evaluation of concept maps is the obstacle for the application of concept maps in instructional situation.

Researchers try to let the computer do the evaluation for humans. Since concept maps include nodes and links, both the relationships and the content of the nodes are important. Kuo-En, Yao-Ting, Rey-Bin, and Shui-Cheng (2005) proposed a weighted concept, whose propositions are given a weight value from 0 to 1. The higher weighted value a proposition is assigned, the more important the proposition is. By comparing a student's map with a teacher's map, each node in both maps gets a closeness index. Then based on the closeness index and weighted value of each node, a score of similarity index is calculated for each node. Using the similarity index, learner's comprehension of the node can be ranked into learned, partially learned, or misconception (Kuo-En et al., 2005).

The limitation in this method is that the nodes in the concept maps are predefined. Students use predefined concepts and links to construct their maps, so the content of the nodes and links are not considered as a factor in comparison. The maps in this paper are open – ended maps, so the algorithms of closeness index and similarity index are not applicable. However, the weighting mechanism is helpful to identify the importance of nodes in expert's map.

Melnik, Garcia-Molina, and Rahm (2002) presented a method called Similarity Flooding Algorithm (SFA). SFA takes two graphs as input, and matches the nodes in both maps. The similarity relies on the intuition that nodes from two maps are similar when their neighbor nodes are similar. The output is a list of corresponding nodes with a similarity value. After a filter selection, the most optimized pairs are considered as the best matched nodes (Melnik et al., 2002). SFA works on many types of graphs, such as data schemas, catalogs, xml ontologies, and concept maps. Particularly, SFA support open-ended nodes, which are required for the project in this paper. This algorithm works for directed labeled graphs, like a process map, which has arrows on links between nodes to indicate the order of the steps of a diagnosis strategy. Melnik et al. (2002) evaluated the accuracy of SFA and conclude that overall labor saving are above 50%, and actual savings might be higher.

Simpson and Dao (2010) present an approach of string comparison with the meaning of the words – semantic similarity. The approach uses WordNet as the database for synonyms. Simpson and Dao (2010) use five steps to compute a semantic similarity for two sentences. The steps are (a) separating sentence into a list of tokens, (b) disambiguating part-of-speech, (c) stemming words, (d) finding the most appropriate sense, and (e) computing the similarity (Simpson, & Dao, 2010). Although Simpson and Dao (2010) node there might be many limitations in the method and need to be improved, their method works fine for the problem in this paper, because of the target learners in this project are trained to use terms in their process maps.

Based on the weighting mechanism, SFA, and semantic similarity of two strings, a combination of these three approaches is adopted in this paper.

Process of Comparison

A process map contains nodes and links between nodes. To compare two process maps, it needs to consider both the relations (links) between nodes and the content of the nodes. We use SFA to match the nodes based on their relationships. The SFA (Fig. 1) represents two input maps semantically in code first, then creates an initial map for the product of each nodes in both maps, calculates their similarities based on the links, and finally generates a list of best paired nodes according to the similarity of each pair (Melnik et al., 2002). During the comparison, WordNet®[1]-based semantic similarity measurement (WSSM) is used to measure the content of nodes.

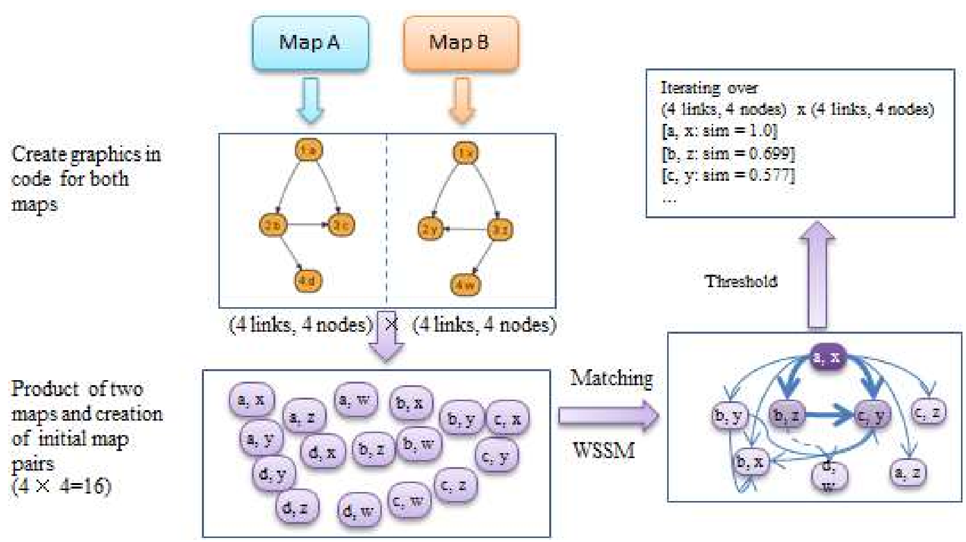

Figure 1: The process of SFA.

[1] WordNet® is a database of English, an open source project based at Princeton University. http://wordnet.princeton.edu/wordnet/

A prototype computer program of process map matching program has been prepared. The process for comparing process maps is shown in Fig. 2. The first step in the prototype program is to generate a base similarity that is calculated based on the result of comparing an expert's map with the expert's map itself. The base similarity includes absolute similarity (s_{aee}) for each paired nodes, according to their links and content. The second step is to compare a novice's map with the expert's map. The results of comparison include absolute similarity (s_{ase}) for each node, which is considered as a matched pair with one node in the expert's map. The third step is to calculate the relative similarity (s_{rse}) for each paired nodes by the percentage of s_{ase} and s_{aee}. And, the overall similarity of the map is calculated based on the relative similarity of each pair. In this step, weighting (w) of each pair can be considered. Important nodes can have more weighting in overall similarity.

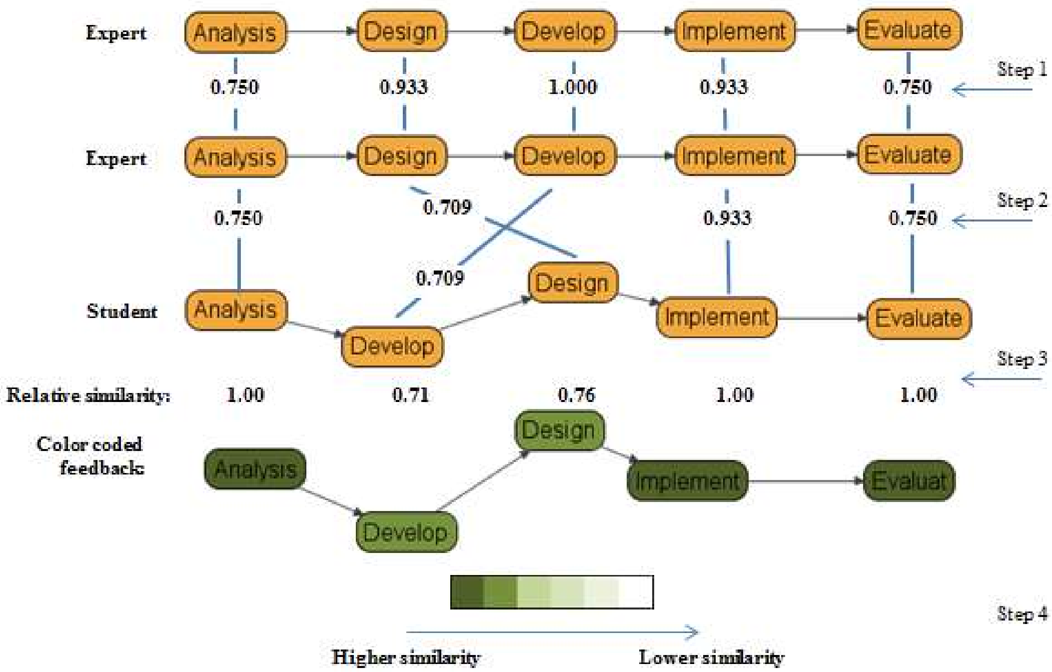

Figure 2: The process of comparison for process maps.

The fourth step is to generate feedback. The feedback includes two parts. One part is a summary that shows the overall similarity, number of nodes in expert's map, number of nodes in learner's map, percentage of matched nodes, and the similarity range of matched nodes, the other part is a color-coded similarity, darker color stands for higher similarity and lighter color stands for lower similarity. A new map is generated based on the learner's map. Sometimes, learner's map doesn't have a high overall similarity because the learner's map has fewer nodes than the expert's, but the feedback still could be positive because the similarities of matched pairs are strong.

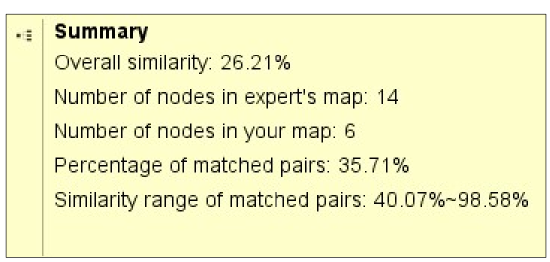

Figure 3: The summary feedback

Fig. 3 is an example of a summary feedback. The overall similarity is 26.21%, because the learner's map contains only 6 nodes and the expert's map has 14 nodes. However, the similarity range of matched nodes is from 40.07% to 98.58%. That means the nodes in the learner's map have high quality as to the nodes in expert's map. When learners get feedback, they can review and analyze which part of their maps might need to be adjusted.

Results of Small Group Test

A formative evaluation (a small group of three try-out learners) for the validity of the comparison method was conducted. For the small group test, an expert created a process map for a problem on computer power supply. Try-out learners were presented with a problem - "A microcomputer fails to turn on", and the task is to develop a visual process map of the steps to diagnose the cause(s) of the problem. All try-out learners were undergraduate students in the majors related to technology and engineering and had the prior knowledge of how a computer works. The try-out learners used the concept map authoring software, VUE[2], to create their process map to represent their approach or strategies to diagnose the power supply problem. The results of the formative evaluation are shown in Tab. 2.

User No.	Overall similarity	Number of nodes in the learner's map	Percentage of matched nodes	Similarity range of match pairs
#1	38%	6	43%	63% ~ 100%
#2	27%	6	43%	40% ~ 88%
#3	14%	5	29%	10% ~ 63%

Table 1: The results of formative evaluation

Comparing the results with experts' opinions, results of user #2 were valid and results of user #3 were reasonable. Results of user #1 were not valid, because user #1's content was not relevant to the target problem.

Discussions

Number of Nodes in a Learner's Map

Based on the formative evaluation, we found that the minimum number of connected nodes must be greater than 4. The result of user #1's map is not valid because the content is not about the target question, but it still got 38% similarity overall and 43% to 100% for match pairs. The reason is that the user didn't connect two nodes correctly in the map. Therefore the map only has 4 connected nodes. The similarity value for each of the four nodes is above 95%, even with totally different content. One possible explanation is that when a small map compared with a large map, the small map will be considered to be included in the large map. In this case, the minimum number of connected nodes must be checked before comparing the map. The program would warn users that the connected nodes in their maps are fewer than 5 in this case, or a percentage based on the number of nodes in expert's map. If the users decided to continue anyway, the program would discard the similarity of the structure and provide the similarity on the content only.

Threshold Values

There are two threshold values in the comparison algorithm. One is *similarity threshold* for the similarity based on the connections (links) of the nodes; the other is *synonym threshold* for the semantic similarity of content. If the similarity is lower than the threshold, it could be considered to be discarded. For example, if the similarity threshold is .25 and the similarity of a pair of nodes is .2, then the two nodes are considered as not similar at all. In the formative evaluation, the similarity threshold was .0, and the synonym threshold was .75.

After the formative evaluation, we tried to adjust the threshold values for the same maps. The similarity threshold was changed to .25, and the synonym threshold was changed to .90. Based on the content of nodes, this pair of values worked better than the values in the formative evaluation, because more unrelated nodes were excluded. Fig. 4 is the summary of the overall similarity values changed before and after adjusting threshold values. The result of user #1, which was not valid before adjustment, becomes reasonable; the overall similarity of user #2

[2] VUE is a visual map authoring tool, an open source project based at Tufts University. http://vue.tufts.edu/

and user #3 are just going down slightly and won't affect original result. To test the validity of the threshold values, more comparison data are needed.

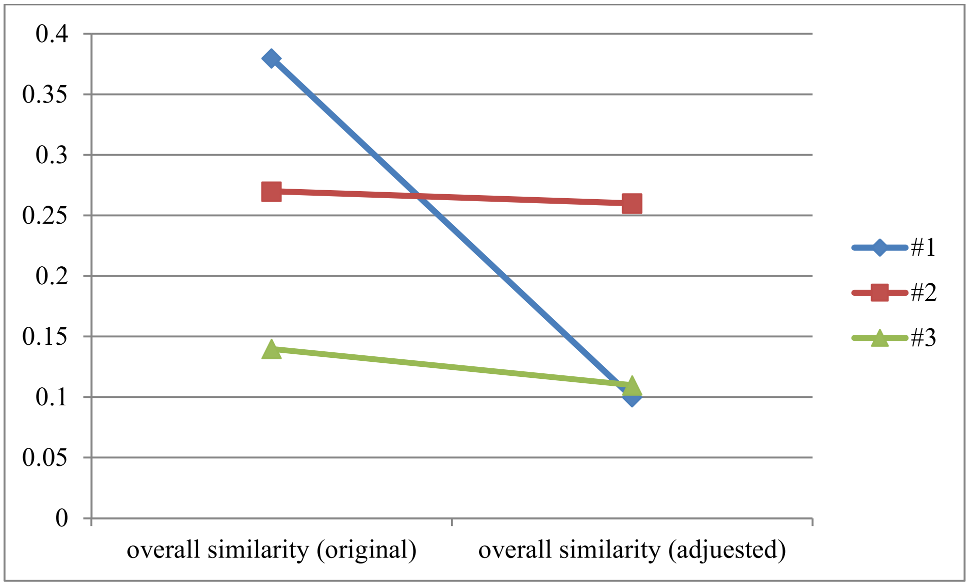

Figure 4: The overall similarity values before and after adjusted threshold values

Correctness of Connections in Map

Maps for people could be drawn approximately. For example, a line between two nodes without actually connected to the nodes is understandable for a human being, because it's obvious for human that the line is connecting the nodes. However, to a computer, the line has to connect the nodes semantically, which means the map creator must connect the nodes by dragging the ends of the line on to target nodes in VUE. Otherwise, the program would omit the link and make the nodes as unconnected. To ensure users can make correct maps, more instruction on how to use VUE to create process map are need; and more validating works in the program are needed as well.

Feedback for Learners

The feedback used in the formative evaluation includes a similarity summary and a color-coded indicator, but only similarity is not enough to help students improve their diagnosis skills. There are two possible improvements for the feedback. The first one is to provide information on the reason of the lower similarity. The similarity of a node contains two score, the relationship with other nodes and the content of the node. Lower similarity might due to lower relationship score or content score. Given these two scores separately, learners could have more guided directions to improve their strategies. The second one is to provide information on the weighting of nodes. Important nodes have larger weighting value. If those important nodes are not presented in learners' maps, it's very possibly to get a lower overall similarity. Given information on what important nodes are missed in learners' maps could help learners reconsider their approaches. For example, the root cause of the power supply problem is the power source. The first three nodes in student #3's map do not hit the root cause of the power source. In this case, if feedback provided suggestions for reconsidering the beginning of the process, the learner might have more chances to improve the original diagnosis approaches.

Conclusions

Visual map is widely used in instructional activities, but the evaluation of visual map is time consuming and instance feedback is not possible through manually evaluation. This paper presents a possible solution by integrating SFA and WSSM for automatic map comparison and providing instant feedback for students. It provides insight into one possible application of visual maps.   The task of comparing two process maps is challenge. It is difficult to gain a high accuracy score or degree on how much two maps are similar automatically by a computer program. Limitations in machine learning and human language understanding might be problems. One word could have different meanings in different situations and one line between nodes could represent contrast ideas. Therefore, the approach presented in this paper is not a perfect one. Because of the number of subjects in the formative evaluation, besides the algorithms, we only found that numbers of nodes, correct connections, and threshold values are factors that affect the results. To provide persuasive evidence, more tests are need in the future.

References

Davies, M. (2011). Concept mapping, mind mapping and argument mapping: what are the differences and do they matter? *Higher Education, 62*(3), 279-301.

Jacobs-Lawson, J. M., & Hershey, D. A. (2002). Concept Maps As an Assessment Tool in Psychology Courses. *Teaching of Psychology, 29*(1), 25-29.

Kassab, S. E., & Hussain, S. (2010). Concept mapping assessment in a problem-based medical curriculum. *Medical Teacher, 32*(11), 926-931. doi: 10.3109/0142159X.2010.497824

Keppens, J., & Hay, D. (2008). Concept map assessment for teaching computer programming. *Computer Science Education, 18*(1), 31-42. doi: 10.1080/08993400701864880

Kuo-En, C., Yao-Ting, S., Rey-Bin, C., & Shui-Cheng, L. (2005). A New Assessment for Computer-based Concept Mapping. *Journal of Educational Technology & Society, 8*(3), 138-148.

Melnik, S., Garcia-Molina, H., & Rahm, E. (2002, 2002). *Similarity flooding: a versatile graph matching algorithm and its application to schema matching.* Paper presented at the Data Engineering, 2002. Proceedings. 18th International Conference on.

Simpson, T., & Dao, T. (2010, Feburary 8). *WordNet-based semantic similarity measurement*. Retrieved from Code Project: http://www.codeproject.com/Articles/11835/WordNet-based-semantic-similarity-measurement

Williams, M. (2004). Concept mapping -- a strategy for assessment. *Nursing Standard, 19*(9), 33-38.

Acknowledgement

Partial support for this work was provided by the National Science Foundation's Transforming Undergraduate Education in Science, Technology, Engineering and Mathematics (TUES) program under Award No. 1140677.