Using Modern Graph Analysis Techniques on Mind Maps to Help Quantify Learning

Peter Jamieson
Deptartment of Electrical and Computer Engineering
Miami University
Email: jamiespa@muohio.edu

***Abstract*—In this work, we use a graph analysis tool to measure how student-created mind maps reflect learning. Mind maps consist of words and connections between words, and this visual tool helps illustrate how an individual understands how these words connect together in a field. From an analysis standpoint, mind maps are graphs consisting of nodes connected by edges. In the fall of 2011, students created three mind maps over the duration of a digital system design course, and at each of the three intervals, these mind maps were created with the same 20 terms that were introduced throughout the course. Each student's mind maps were then digitally encoded and analyzed using a modern graph analysis tool called GraphCrunch II. Our results show that a simple analysis of graph density is a poor indicator of learning since this metric does not capture a graph's structure, and it is this structure that reflects meaning and understanding by the learner. Instead, a metric called relative graphlet frequency distance (RGF-distance), which is calculated by comparing a golden mind map (expert created mind map) to each of the students mind maps, is used to analyze each students understanding of how these words relate. Our results show that learner's mind maps decrease in RGF-distance over the period of the course, and this means that the students are building graphs more similar to that of the golden model. We, also, see that the RGF-distance over the set of students compared to their grades on an exam or overall grade in the course has some correlation, meaning that these mind maps relate to grades in terms of the learners understanding of vocabulary, but the correlation is not strong. The ultimate goal of these tools is to provide the learners with a method of getting automatic feedback on their understanding as well as learning progress in particular topics.**

I. Introduction

Concept maps [1], [2] and mind maps [3] are useful tools in education, and they can be used both in lectures as classroom assessment techniques (CATs) and outside the classroom for the learner to express ideas in a visual form. Mind maps are words/concepts connected by lines where a line indicates a relationship between the words, and concept maps (also called knowledge maps) have similar structure except the connecting lines have directionality (indicated by an arrows) and have associated prepositional phrases that indicate how the words are related. For this work, maps are mathematically known as graphs, and therefore, we can leverage research in the understanding of graphs to help analyze mind maps.

Our focus is on using modern graph analysis tools to evaluate in class, student-created mind maps. Our goal is to help build tools that learners can use with their mind maps to automatically give them feedback on their understanding of a topic. To evaluate a map, either the map can be evaluated by an expert observer, or the map can be compared to a golden map generated by an expert. The first of these evaluation options is very difficult for computers to perform since the computer needs some inherent understanding of the topic itself. The later is the approach used since a machine can use the golden map and compare it to the learners map without any knowledge of the subject.

To compare learner mind maps to golden maps we use a tool called GraphCrunch II [4], which is a tool developed for analyzing the similarities between proteins. This tool uses a unique method to compare graphs and is easily applicable to mind maps. Over the 2011 semester of a course on digital design, we collected student mind maps (course beginning, post exam I, and post exam II). These mind maps were CATs where students had 10 minutes to create their maps that consisted of 20 terms, and these 20 terms were introduced throughout the semester. Each student's mind maps was then digitally encoded and analyzed using GraphCrunch II. Our results show that another metric, graph density (defined later), is a poor indicator of learning since this metric does not capture the graph's structure, and it is structure that reflects meaning and understanding by the learner.

Instead, a metric called relative graphlet frequency distance (RGF-distance) [5] is calculated by GraphCrunch II. Our results, with this metric, show that the majority of learners have a decrease in RGF-distance for their maps over the course, and this means that the students are building graphs more similar to the golden map. We consider this result evidence of student's learning the vocabulary of a field and the relationships between words. We, also, see that the RGF-distance for students compared to their grades on an exam or overall grade in the course has some correlation, meaning that these mind maps reflect performance in the course. However, the correlation coefficient is not strong enough to consider using the graph analysis tool as an assessment measure.

The remainder of this paper is organized as follows. Section II reviews previous work in analyzing mind maps and introduces more details about graphs and the GraphCrunch II tool to help the reader. Section III describes the experiment, and Section IV shows our results. Finally, Section V concludes the paper.

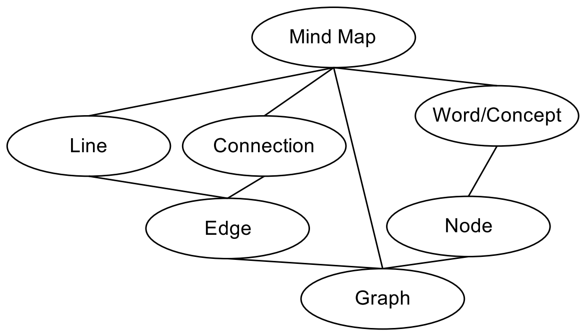

Fig. 1. Example of a mind map on the relationship between mind maps and graphs

II. Background

Mind maps [3] are a visual tool that can be used to represent the connections between a number of concepts. These maps are useful CATs [6], and both the student and teacher can observe how concepts connect in the learners mind, which allows teachers to provide feedback. Figure 1 shows an example mind map that expresses the author's understanding of mind maps and how they relate to mathematical graphs. Basically, the words/concepts that are in a mind map are the nodes of a graph (circled bubbles), and the connecting lines between these words are edges of a graph. Concept maps [1], [2], also called knowledge maps, are more complex than mind maps since the edges are labeled with prepositional phrases. We do not analyze concept maps in this work, but they are also powerful learning tools with a valuable research literature that compliments and motivates this work. For a good introduction to concept maps, take a look at Novak's book, *Learning, Creating, and Using Knowledge: Concept Maps As Facilitative Tools in Schools and Corporations* [7].

A. Map Analysis Classifications

As described earlier, our focus is on automatically analyzing mind maps using modern graph analysis tools. To do this, we will compare student maps to a golden model; Ruiz *et. al.* [8] called this type of scoring *comparison with a criterion map*. Ruiz identified another scoring mechanism called *score map*, which includes basic graph analysis techniques such as counting edges and nodes, and they describe a third option, *hybrid model*, that combines the two. Their work looks at five other studies with concept maps and categorizes them based on their scoring technique. In this work, our approach is the third choice, hybrid scoring, since we use the criterion map to assist the automated scoring.

Herl *et. al.* [9] further categorize map analysis based on what the learner is allowed to do. They call maps that are restricted in construction as *closed*, which means the words and concepts are limited, and *open* maps are unrestricted. A number of researchers [10], [2], [11], have presented ways of defining closed maps for concept maps, and in our work, the mind maps are closed since we restrict the number of words/concepts that a learner can include in their maps.

B. Automated Map Analysis

As a precursor to automated map analysis, various methods of scoring a map have been proposed by researches, and a brief list of scoring metrics in the literature include the following:

- Concept map - count the number of valid propositions, levels of hierarchy, examples, and crosslinks [1] where weights can be introduced to each count ([12], [13])
- Concept map - a measure of *hierarchiness* which relates hierarchy in the map [8]
- Mind map - compare the scores on tests to the technique [14]
- Concept map - the more important a concept, the closer it is to the top of the tree [15]
- Mind map - have two independent experts score (sometimes with a rubric) the mind map on a scale two times with one week delay and compare correlation of ratings [16]

The scoring of maps has been challenged by many researchers in the literature [17], [18]. Kinchin and Hay [19] criticize the shortcomings of strict scoring of concept maps as a motivation to propose qualitative analysis of maps. Interestingly, one of their key contributions in this qualitative approach is looking for *spoke*, *net*, and *chain* like structures in a map, and these structures are internal graph structures that are captured by what will later be described as graphlets.

Early attempts at automating the analysis of maps and providing feedback focused on hint like mechanisms. A criterion map (called scaffold in these works) provides student creators with hints for their maps on what is missing and what does not belong [20], [21]. Conlon [22] built a system that used Novak's scoring mechanism and other artificial intelligent concepts to build an open concept map creation system that provided feedback to students. Our goal is study modern graph analysis tools for similar purposes.

C. Graph Analysis

We have already introduced concepts of *nodes* and *edges*. We should also mention that mind maps are classified as undirected graphs, which means that the edges do not have a direction, normally indicated by an arrow; concept maps, on the other hand, are directed graphs. In the rest of this section, we will provide a better understanding of the metrics compared as our experimental measures - density of a graph and relative graphlet frequency distance (RGF-distance). Note that mathematical formulas and definitions are provided in graph notation, but the reader does not need to understand this notation to understand this work.

The *density* of a graph is defined as the total number of edges ($|E|$) divided by the maximum number of edges that a graph could have ($.5*|V|*(|V|-1)$). For the graph in figure 1, the maximum number of possible edges is 21 and the number of edges is 9, so the density is 9/21 or, approximately, 0.43. The greater this value, to a maximum of 1, means the more connectivity between the nodes within a graph.

Density measures how much connectivity there is in a graph. In terms of relating this to mind maps, we might hypothesize that the more connected a mind map is then the more tightly related the topics are. If we compare metrics for a student map to the criterion map then we might say the closer these metrics are to one another than the global connectivity is more similar. The problem, however, is the structure of the mind map is not captured with simple metrics and students might be making wrong connections that somehow make the comparative metrics more equal.

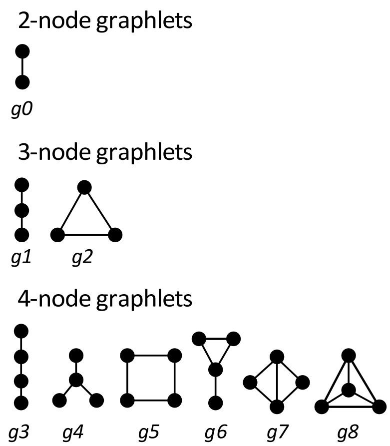

Fig. 2. Graphlets of size 2, 3, and 4

Before defining RGF-distance, we first introduce graphlets. Graphlets, formally, are "a connected network with a small number of nodes" [5] and these small graphs are non-isomorphic induced subgraphs of a larger graph. Figure 2 shows *all* the graphlets of size 2, 3, and 4. Note that the graphlet of size 1 is a single unconnected node and is not that useful.

The power of the graphlet is how it can be used to analyze a graph. The procedure developed by Przulj *et. al.* [5] is to search for all graphlets of size 3, 4, and 5 in a given graph. Based on the count of each type of graphlet, we then can construct a signature in the form (g1, g2, g3, g4, g5, g6, g7, g8, ..., g28, g29), where g1 is number of the first type of graphlet of size 3 shown in figure 2 and g29 is the count for the last graphlet of size 5. This signature can be compared to another graphs signature to get a measure of similarity, and Przulj *et. al.* used their technique to compare graphs representing biological structures such as proteins.

RGF-distance is a measure of the difference in frequency of graphlets of g1, g2, g3, ..., g28, and g29 appearing in the two graphs being compared. A detailed equation is presented in Przulj *et. al.* [5] and the reader can find the details on the calculation of RGF-distance. For this work, we must understand that GraphCrunch II will calculate this metric for us, and the smaller this number gets means the more similar the two graphs are.

RGF-distance captures the comparative structure of two graphs, and in terms of comparing student mind maps to the criterion map, we hypothesize that the lower the RGF-distance is means that students better understand the relationships between concepts/words since their maps have more similar structure to the expert's criterion map. This is the main hypothesis of this work.

III. Semester Mind Map Experimental Setup

For our work, the goal is to automatically analyze student-created mind maps over a semester long course and observe how these students are learning the course vocabulary as reflected by their mind maps. The focus course for this experiment is a digital design course offered at the 200 level. The course starts with how transistors can be organized to make basic Boolean gates and ends at designing finite state machines using a hardware description language (HDL). From our perspective, the most challenging aspect for most students is the application of HDLs to design hardware as the language differs significantly from sequential programming languages that students are much more familiar with. However, mind maps only play a small part in understanding HDL application, and this work focuses on the students understanding of digital system vocabulary.

Table I shows a summary of our experiment. Column one describes the category, and column two and three lists the category type and specific details.

Our mind map experiments are on closed mind maps. Specifically, a list of 20 terms are provided that are introduced in the course. The following list of the 20 terms is ordered chronologically based on when the word/concept is introduced in the course. Note that by exam I, terms 1 through 15 have been introduced, and by exam II the remaining words have been presented. Also, note that the list order is randomly presented to the students. The list includes the following: Electricity, Transistor, Vdd, False, Digital, AND, Truth table, XOR, Schematic, Power, Timing, Binary, Decimal, Two's compliment, Multiplexer, Sequential, Register, HDL, Always, FSM.

During the second class of the course, we introduce mind maps using an illustration of constructing mind maps for countries. We show how the mind map can be constructed differently depending on if we are thinking about geographical location, oil supply, or military alliances and enemies. After this basic training, we then show the list of 20 terms and give the students 10 minutes to create their first mind map. This is repeated after exam I and exam II with the same terms and the same amount of time. This means for each student who has chosen to participate (our IRB approved protocol allowed students to remove their participation agreement any time in the semester before final marks were released) could have created up to three mind maps over the semester.

To use these paper-based mind maps with graph analysis tools such as GraphCrunch II, the mind maps need to be digitally encoded. One of the digital formats that GraphCrunch uses is a simple line-by-line graph representation where two words are included on each line; the pairing of words on a line means that there is an edge between them. We built a simple key-press data entry program based on the first letters of each word to quickly encode the paper mind maps into a

TABLE I
A SUMMARY OF THE PROPERTIES OF OUR MIND MAP EXPERIMENT

Evaluation Category	This Experiment	Details
Mind map option	Closed	Restricted to 20 terms
Evaluation	Hybrid	Criterion map that is used for scoring
Scoring metric	Various	Comparison of simple graph metrics to RGF-distance
Mind map creation	In class	10 minute activity at the start of class

digital form. Each student is given a numerical id, and each of their mind maps is encoded and a corresponding grade entry is labeled with their id. In this way, students become anonymous once all their data is in as per our IRB protocol. With the digitally encoded mind maps, the graphs are analyzed.

IV. EXPERIMENTAL RESULTS

The hypothesis is that the RGF-distance metric, which is a comparison of graph structure between student and the criterion mind maps, will decrease in value over time and show that student's mind maps reflect a better understanding of how the words and concepts in a field are organized. This learning will be examined in two ways; first, we expect that there will be some correlation between grades and RGF-distance, and second, over the course, we expect an individual's mind maps to show a better understanding of how the words/terms in a field relate. In addition to these two trends, we will first look at how simple graph metrics relate to RGF-distance assuming that these metrics can not capture structure that RGF-distance can.

A. Comparison of RGF-distance to other Graph Metrics

We speculate that the RGF-distance metric is a better metric for evaluating mind maps since it captures comparative graph structure. Here, we compare graph density to the RGF-distance metric in two ways to show that RGF-distance is more useful than simple metrics. First, for each set of mind maps the correlation coefficient is compared to the respective RGF-distance, and second, the metrics for the criterion model are compared to see if there is a relationship.

TABLE II
STATISTICS FOR THE GRAPH METRICS ON MIND MAPS

Correlation - Density to RGF-distance			Golden Model Stats		
Course Start	Exam I	Exam II	Edges	Nodes	Density
-0.14	-0.47	-0.93	47	20	0.25

Table II summarizes this data. The first three columns show the correlation coefficients, and the last three columns show stats on the criterion mind map. In terms of the correlation coefficients between density and RGF-distance there appears to be none for the first two mind maps. However, for the post exam II mind maps, there is a high correlation. To understand if the correlation has any meaning, we look at the second aspect of analyzing this data, how does the density compare to the densities in the student mind maps. On average the density for the three sets of mind maps (0.16, 0.15, and 0.14 respectively) are all smaller than the density for the criterion map (0.25), which suggests the similarity is small. The best mind maps (as a measure of RGF-distances below 5) created in the post exam II do have the higher densities and are similar to the criterion maps density. This is not the case for the mind maps created at course start and post exam I, so we have to conclude that graph density could be a quick indicator of similarity that students could use as they create their mind maps, but by itself it is not sufficient to help students.

B. RGF-distance Relation to Grades

To examine how RGF-distance for student generated mind maps compares to scores on class exams, we will plot RGF-distance on x-axis and the exam or final grade score out of 100 on the y-axis. We will then plot a logarithmic regression line on the graph to get some understanding of the general relationship between grade and RGF-distance. Finally, we will calculate correlation coefficients between grade and RGF-distance for each time the mind maps were created to evaluate if there is a strong linear relationship.

Figure 3 shows these scatter plots. Starting with the graph in the upper left, the blue points are the mind map RGF-distance from the start of the course (n=40) compared to the student's grade, and the red points are the same but for the mind maps created post exam I (n=24).

One interesting observation is there is a shift of RGF-distance from the start of class to the first exam, which signifies that the mind maps are more similar to the criterion map post exam I. The trend line, in this first graph, suggests that there is some relationship between grade and RGF-distance, but the actual correlation coefficients are weak, -0.14 and -0.27 respectively. The second two graphs, which compare the mind maps post exam I and exam II (n=24), do not show a shift in RGF-distance between the two periods of time. This is probably due to the fact that the majority of vocabulary is introduced by the first exam and the shift would only be minor. We see similar trend lines to the first graph, and there is a high correlation coefficient for mind maps created after the first exam to the grade (-0.45 for exam II and -0.63 for final grade). However, since the number of participants and the specific participants at each stage (40 students for the class start mind map, 24 students for the post exam I mind map, and 24 students for the post exam II mind map) is different, these correlation and trend lines are only a hint of the relationship between mind map structure and grade. Interestingly, the correlation for the final mind maps RGF-distance to grades is not as high as post exam I (-0.31 for

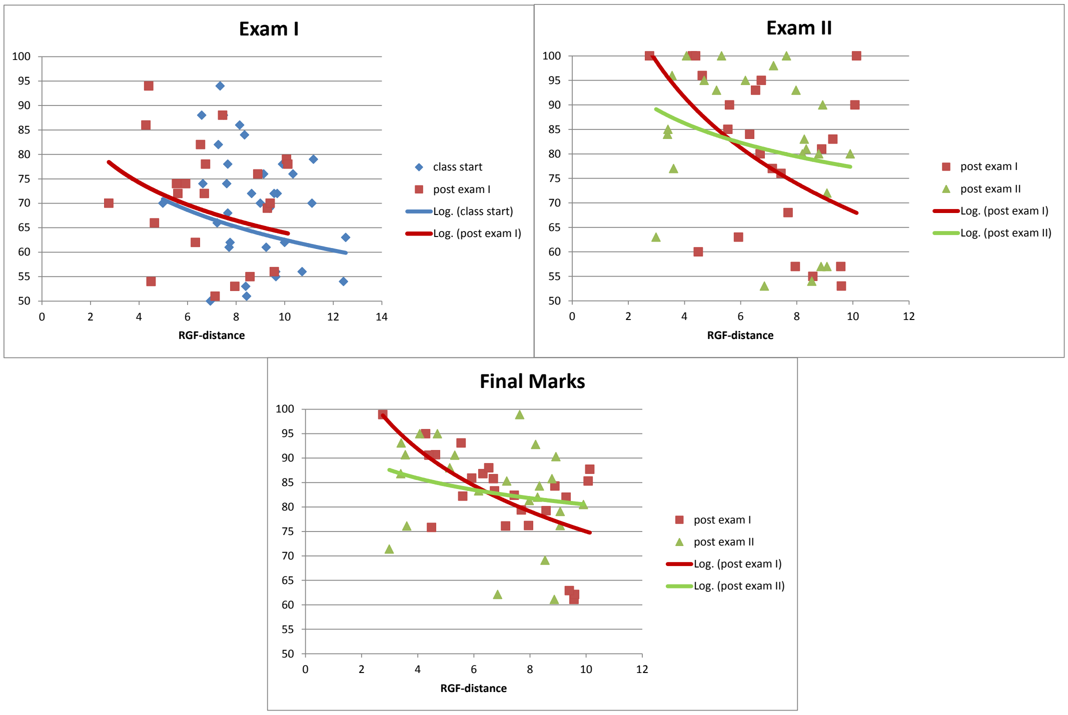

Fig. 3. Scatter plots for grade to RGF-distance on exams and overall grade with RGF-distance on the x-axis and grade point out of 100 on the y-axis.

exam II and -0.26 for final grade).

The greatest correlation is between final grades and post exam I mind maps with a correlation coefficient of -0.63. We hypothesize that mind maps RGF-distance compared to overall grade is a better comparison since the mind map is an indicator of understanding on how concepts connect, and this type of knowledge is more important to activities in the overall course, which is assessed on exams, labs, and projects. The exams, though needing an understanding of vocabulary, is more focused on testing students problem solving and design skills, which is not a good assessment of understanding how concept/words are related.

C. *Individual RGF-distance Analysis*

Figure 4 shows four graphs where each graph is a bin of students with a certain final grade range. Each graph shows the measure of RGF-distance (y-axis) for the students mind maps over the semester (x-axis); in the graph a line connects two points if they are for the same student and they have created mind maps at two adjacent times. Looking at all the graphs, we can see that for the most part, RGF-distance measurements decrease, meaning the student mind maps are becoming more similar in structure to the criterion map. We believe this is strong evidence supporting our hypothesis. The outliers to this trend tend to be found when comparing the RGF-distance measure from mind maps created post exam I and post exam II. Whether this reflects student confusion, poor performance on the mind map activity, or some other factor is unknown, but these trends are observable in each of the grade bins, and therefore, is not related to students overall grade performance in the course.

V. Conclusion

In this work, we investigated the viability of using RGF-distance as a graph metric that could be used to automatically give students feedback on their own mind maps. We compared other simple graph metrics to verify that RGF-distance is a superior metric based on how it compares graph structure. Next, our results showed that there is some relationship between final grades and RGF-distance measured for student made mind maps compared to a criterion model. This relationship, however, was not strong, and we do not believe that mind maps could be used to assess students on activities that are not vocabulary focused. When understanding a field's vocabulary is the focus, the mind map evaluated in terms of RGF-distance could be used as a grading tool, but further exploration is needed in this area.

Our analysis of individual performance of a student over the term showed that mind maps, in general, improve over time based on a decrease in RGF-distance. This result is exciting since we could imagine a scenario where a student could use a tool over the term to see how their understanding of topic vocabulary has improved compared to their previous attempts. This type of scoring mechanism is useful to provide an overall evaluation of the mind map, but it does not provide a mechanism to provide detailed feedback to the learners.

References

[1] J. D. Novak and D. B. Gowin, *Learning how to learn*. New York: Cambridge University Press, 1984.

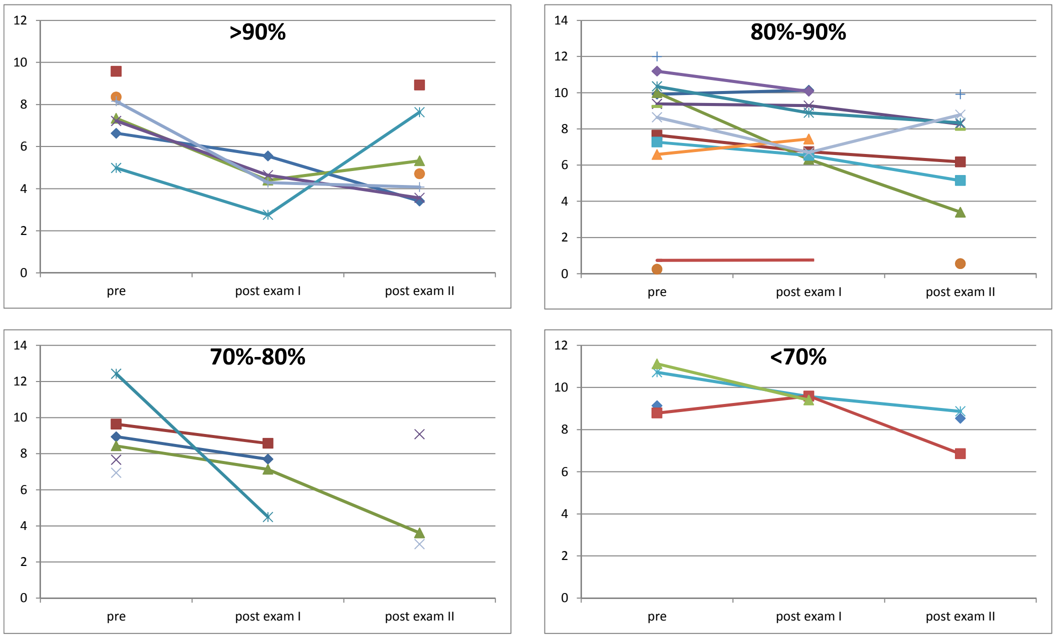

Fig. 4. Line graphs for student mind maps RGF-distance over semester sorted into final grade bins with time of mind map on the x-axis and RGF-distance metric on the y-axis.

[2] D. F. Dansereau and C. D. Holley, *Development and Evaluation of a Text Mapping Strategy*. Amsterdam, New York & Oxford: North-Holland Publishing Company, 1982, pp. 536–554.

[3] T. Anderson, *Study skills and learning strategies*. Center for the Study of Reading, University of Illinois at Urbana-Champaign, 1978. [Online]. Available: http://books.google.com/books?id=JNHUjwEACAAJ

[4] O. Kuchaiev, A. Stevanovic, W. Hayes, and N. Przulj, "Graphcruch 2: Software tool for network modeling, alignment and clustering," *BMC Bioinformatics*, vol. 12, p. 24, 2011.

[5] N. Przulj, D. G. Corneil, and I. Jurisica, "Modeling interactome: scale-free or geometric?" *Bioinformatics*, vol. 20, no. 18, pp. 3508–3515, 2004.

[6] T. Angelo and K. Kross, *Classroom Assessment Techniques - A Handbook for College Teachers*. Jossey-Bass, 2003.

[7] J. D. Novak, *Learning, Creating, and Using Knowledge: Concept Maps As Facilitative Tools in Schools and Corporations*. Lawrence Erlbaum Assoc Inc, 1998.

[8] M. Ruiz-Primo, R. Shavelson, and S. Schultz, "On the validity of concept-map-based assessment interpretations: An experimental testing the assumption of hierarchical concept maps in science," University of California, Tech. Rep., 1997. [Online]. Available: http://research.cse.ucla.edu/Reports/TECH455.PDF

[9] H. E. Herl, H. F. O'Neil, G. K. W. K. Chung, and J. Schacter, "Reliability and validity of a computer-based knowledge mapping system to measure content understanding," *Computers in Human Behavior*, vol. 15, no. 3-4, pp. 315–333, May 1999. [Online]. Available: http://dx.doi.org/10.1016/S0747-5632(99)00026-6

[10] A. M. Collins and M. R. Quillian, *How to Make a Language User*. Academic Press, 1972, pp. 309–351.

[11] J. Lambiotte, D. Dansereau, D. Cross, and S. Reynolds, "Multirelational semantic maps," *Educational Psychology Review*, vol. 1, pp. 331–367, 1989. [Online]. Available: http://dx.doi.org/10.1007/BF01320098

[12] T. Stoddart, R. Abrams, E. Gasper, and D. Canaday, "Concept maps as assessment in science inquiry learning - a report of methodology," *International Journal of Science Education*, vol. 22, no. 12, pp. 1221–1246, 2000. [Online]. Available: http://www.informaworld.com/openurl?genre=article&doi=10.1080/095006900750036235&magic=crossref

[13] D. C. West, J. K. Park, J. R. Pomeroy, and J. Sandoval, "Concept mapping assessment in medical education: a comparison of two scoring systems." *Medical Education*, vol. 36, no. 9, pp. 820–826, 2002. [Online]. Available: http://www.ncbi.nlm.nih.gov/pubmed/12354244

[14] I. Abi-El-Mona and F. Adb-El-Khalick, "The influence of mind mapping on eighth graders science achievement," *School Science and Mathematics*, vol. 108, no. 7, pp. 298–312, 2008. [Online]. Available: http://dx.doi.org/10.1111/j.1949-8594.2008.tb17843.x

[15] D. Leake, A. Maguitman, and T. Reichherzer, "Understanding knowledge models: Modeling assessment of concept importance in concept maps," in *In Proceedings of CogSci 2004. Mahwah, NJ: Erlbaum. In*, 2004, pp. 785–800.

[16] E. Evrekli, D. Inel, and A. Galim, "Development of a scoring system to assess mind maps," *Procedia - Social and Behavioral Sciences*, vol. 2, no. 2, pp. 2330 – 2334, 2010. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S187704281000371X

[17] M. A. Ruiz-Primo and R. J. Shavelson, "Problems and issues in the use of concept maps in science assessment," *Journal of Research in Science Teaching*, vol. 33, no. 6, pp. 569–600, Aug. 1996. [Online]. Available: http://dx.doi.org/10.1002/(SICI)1098-2736(199608)33:6\%3C569::AID-TEA1\%3E3.0.CO;2-M

[18] X. Liu and M. Hinchey, "The internal consistency of a concept mapping scoring scheme and its effect on prediction validity," *International Journal of Science Education*, vol. 18, no. 8, pp. 921–937, 1996. [Online]. Available: http://www.tandfonline.com/doi/abs/10.1080/0950069960180805

[19] I. M. Kinchin, D. B. Hay, and A. Adams, "How a qualitative approach to concept map analysis can be used to aid learning by illustrating patterns of conceptual development," *Journal of Educational Research*, vol. 42, pp. 43–57, 2001. [Online]. Available: http://www.personal.psu.edu/kmo178/blogs/kmorourke/qualitative%20approach%20to%20concept%20map%20analysis.pdf

[20] K. Chang, Y. Sung, and S. Chen, "Learning through computer-based concept mapping with scaffolding aid," *Journal of Computer Assisted Learning*, vol. 17, no. 1, pp. 21–33, 2001. [Online]. Available: http://dx.doi.org/10.1111/j.1365-2729.2001.00156.x

[21] G. Biswas, D. Schwartz, and J. Bransford, "Smart machines in education," K. D. Forbus and P. J. Feltovich, Eds. Cambridge, MA, USA: MIT Press, 2001, ch. Technology support for complex problem solving: from SAD environments to AI, pp. 71–97. [Online]. Available: http://www.vuse.vanderbilt.edu/~biswas/Research/ile/papers/sad01/sad01.pdf

[22] T. Conlon, "But is our concept map any good?: Classroom experiences with the reasonable fallible analyser," in *Proc. of the First Int. Conference on Concept Mapping*, 2004. [Online]. Available: http://cmc.ihmc.us/papers/cmc2004-054.pdf