Scoring Concept Maps: an Overview

Alla Anohina, Janis Grundspenkis

Abstract: *The paper presents summarization of proposed scoring schemes of concept maps. The main attention is devoted not only to theoretical approaches considering assessment of students' concept maps, but also to scoring mechanisms implemented in the existent concept map based computerized assessment systems. Conclusions about application of examined scoring schemes are made and factors affecting students' score in computerized assessment systems are identified.*

Key words: *Concept Map, Concept Map Based Assessment System, Scoring Scheme.*

INTRODUCTION

Obvious necessity to assess the organization of students' knowledge has led to the appearance of various tools of visual representation of knowledge structure. One of such tools is a concept map (CM) which can foster the learning of well-integrated structural knowledge and externalize the conceptual knowledge (both correct and erroneous) that learners hold in a knowledge domain [1]. A CM is a graph with labeled nodes corresponding to concepts in a problem domain and with arcs (directed or undirected and with or without linking phrases) indicating relationships between pairs of concepts. A linking phrase specifies the kind of a relationship between concepts. A semantic unit of a CM is a proposition stated as a concept-link-concept triple and presenting a meaningful statement about some object or event in the problem domain [1].

Development of CM based computerized assessment systems faces a problem: what kind of a scoring scheme for CMs should be chosen from a huge variety of proposed schemes? The scoring method of student outcome maps is a crucial concern when using knowledge mapping in an assessment setting [12]. Most of the assessment schemes proposed in literature either have been applied to studies where the evaluation of CMs is human-based or constitutes a theoretical framework, while the number of systems that have embedded a scheme for automated assessment is minimal [8].

At the Department of Systems Theory and Design of Riga Technical University the development of the adaptive CM based knowledge assessment system (KAS) started in 2005. At the moment the KAS has reached maturity concerning its architecture and operating principles [9, 26]. At the same time developers of the KAS face the same common challenge- elaboration of an adaptive automated scoring scheme for evaluation of students' CMs. The paper presents results of the first step towards solution of the mentioned problem- summary on examined known CM scoring schemes.

The paper is organized as follows. The next section gives an overview of the proposed scoring schemes of CMs. Then scoring schemes implemented in the existent CM based computerized assessment systems are described. Finally, conclusions and directions of future work are presented.

SCORING SCHEMES OF CONCEPT MAPS

According to [22] a scoring scheme is a systematic method with which students' CMs can be evaluated accurately and consistently. Three main approaches are the following: 1) evaluation of components of a CM; 2) comparison with an expert CM, and 3) combination of both. In [8], taking into account that scoring of CMs usually is accomplished by comparing a student's map with an expert's one, it is pointed out that two most commonly investigated assessment methods are the structural method, which quantitatively assesses CMs considering valid components, and the relational method, which focuses on the accuracy of each proposition. On the basis of analysis of literature we propose classification of scoring schemes using five criteria: a) type of scoring: quantitative, qualitative or combination, b) scoring method: structural, relational or combination,

Table 1: Scoring schemes of CMs

Scoring scheme	Characteristics*
Different number of points for valid propositions, levels of hierarchy, important cross-links, not illustrative cross-links, and examples [19]	QN, S, NE, H, R: hierarchical CMs
Improvement of Novak's and Gowin's scheme [19] by granting points for valid branchings [27]	QN, S, NE, H, R: hierarchical CMs
Closeness index calculated taking into account set of concepts in student's and expert's map [7]	QN, S, E, A, NR
Number of links, hierarchies, cross-links and examples [13]	QN, S, NE, H, R: hierarchical CMs
Improvement of Novak's and Gowin's scheme [19] by granting points for the first level branching and for each next level [15]	QN, S, NE, H, R: hierarchical CMs
Stringent semantic content score, categorized semantic content score, organizational structure score, number of terms and links used [10]	QN, S+R, E, A, NR
Different number of points for valid concepts, relationships, the first instance of branching, an additional instance of branching, levels of hierarchy, and cross-links [20]	QN, S, NE, H, R: hierarchical CMs
Validity of propositions, concepts used on the nodes, number of cycles and arrows, a total proposition validity score, a congruence score, a salience score, a beginner-nodes score, a cycle score, a circuit score, a subordinate score [22]	QN, S+R, E, H, NR
Holistic approach (judgment about overall understanding of the concepts assigning a score on a scale from 1 to 10), structural method of Novak and Gowin [19], relational method in which each proposition is scored from 0 to 3 points taking into account such criteria as a label of relationship, direction of the arrow, etc.[16]	QN, S+R, E+NE, H, R: hierarchical CMs
The total number of useful, incorrect, incomplete, emerging and defined links, and different levels of link utility (level 1 - an example, level 2- a fundamental fact, level 3- a link explained by other links) [18]	QN, S, NE, H, NR
Different points for central concepts from an expert map, each of the remaining concepts from an expert map, any of "external" concepts which are not part of the expert map, as well as for expert relationships, other valid and relevant relationships between two expert concepts or between one expert concept and one "external" concept, valid and relevant relationships between two "external" concepts [23]	QN, S+R, E, H, NR
Different points for valid propositions, but which are less like what experts might include in their CMs, valid propositions found in either expert's map, propositions identified as critical in an expert's map, categorization of propositions (illogical, pragmatic, scientific and principled) [12]	QN+QL, R, E, A, NR
Different number of points for valid propositions, cross-links, hierarchies, and examples, for valid relationships with incorrect propositional label, with a correct propositional label, but missing foundational or core relationship to subject matter, and with a correct propositional label and apparent foundational or core relationship [28]	QN, S+R, NE, H, R: hierarchical CMs
Different number of points for wrong or scientifically irrelevant, partially incorrect, correct but scientifically "thin", and scientifically correct and scientifically stated propositions, the sum of individual proposition scores, proposition choices and structure complexity [30]	QN+QL, R, NE, H, NR
Scoring of propositions and calculation of the similarity index as the sum of scores of student's propositions divided by the sum of propositional scores in the expert map [3]	QN, R, E, A, NR
Score based on link lines between concepts and the geometric distances between concepts[24]	QN, R, E, A, R: does not take into account labels of relationships

*QN- quantitative scoring, QL-qualitative scoring, QN+QL – combination; S- structural method, R-relational method, S+R- combination; NE- does not use an expert map; E- usage of an expert map, E+NE – combination; H- human-based scoring, A -automatic scoring; R-restriction, NR- no restrictions

c) scoring on the basis of an expert map or without its usage, d) scheme offered or used for automatic or human-based scoring, and e) presence or lack of restrictions concerning application of a scheme.

Table 1 summarizes a number of CM scoring schemes (ordered in chronological order) and clearly shows that most schemes are based on quantitative measures and only few of them combine both quantitative and qualitative approaches while schemes applying purely qualitative methods are absent.

The structural scoring scheme first proposed by Novak [19] and afterwards modified by several authors [15, 16, 20, 27, 28] mainly is applicable only for hierarchical CMs because such aspects as levels of hierarchy and cross-links are taken into account.

Comparing a student's map with the expert's one a similarity or closeness index is usually calculated that shows the extent to which the CM of a student matches that of an expert. The index can be calculated taking into account only structural components of both maps or also considering validity of propositions. The most widely-known closeness index is the Goldsmith's index [7] comparing only sets of concepts in an expert map and in a student's map. As a result, other similarity index related to propositions is proposed in [3]. Moreover, some researchers have used an approach when a student's map is compared with two or more expert maps. In [12] CMs are scored by comparing each student's CM with two expert maps. In turn, Herl and colleagues [10] use a matching algorithm, which includes several expert maps evaluating each student's map. In [6] populations of CMs semantically comparable to students' CMs are built from a teacher's CM. This allows assessing of the knowledge level of each individual student taking into account that humans can construct the same knowledge differently.

Recently assessment of CMs through consideration of propositions has become a topical research direction. In [17] authors hypothesize, that the distance between concepts is in inverse relation with the level of understanding or knowledge about these concepts links and provide a way how to analyze spatial arrangement of a CM on the basis of normalized distances between concepts. In [14] it is pointed out that there are few analytical schemes for identifying implicit linkage patterns in CMs. The authors combine quantitative assessment based on the closeness index of a particular concept with the analysis of structural linkage patterns and identify confused concepts, substitute concepts and hidden wrong concepts. In [29] a novel scoring algorithm based on proposition chains (a linked list consisting of all the propositions in one of the longest paths in the graph) is proposed. It is intended for use in computer-based concept mapping systems and its purpose is to represent and assess the semantic meanings between propositions through the mode of proposition chain.

According to Table 1 the greatest part of the offered schemes are intended for human-based scoring. Thus, it is very difficult to evaluate, whether they are feasible and useful in CM based knowledge assessment systems.

COMPUTER-BASED CONCEPT MAPPING ASSESSMENT SYSTEMS

Regardless of the fact that at present the number of computer-based concept mapping assessment systems is small, they differ in their complexity. One of the simplest systems is presented in [4]. It allows students to create a map from a given list of concepts and relationships and to receive feedback previously defined by the teacher for different propositions. However, scoring of CMs is not performed.

The other example is the system WCOMT [25] in which a student is provided with a series of test questions presented in the form of a fragment of an incomplete CM containing blanks which must be filled-in. The CM blanks are scored with different weights ranging from 1 to 3 points depending on the importance of a blank. The scoring scheme used in the system fully conforms to the task offered.

A computerized knowledge assessment tool TPL-KATS [11] providing a task of creating a CM from a given list of concepts and relationships uses three basic scoring methods: minimum valence (the minimum strength of the relationship on the shortest path between two concepts), shortest path (the number of links in the shortest path between two concepts), and average valence (is calculated by taking the shortest path between two concepts and averaging the strength values of each link along that pathway).

In turn, more complex systems can provide a range of tasks and complicated scoring schemes. In [2] two versions of the system are described: "construct-by-self" where students construct CMs by themselves and "construct-on-scaffold" in which students are presented with an incomplete expert CM with some blank nodes and links. In both cases lists of concepts and linking phrases are provided. Moreover, the system provides hints on request. The analysis of a student's map is performed by comparing it with the predefined expert CM and calculating the similarity index. For the "construct-by-self" version criteria of Novak and Gowin [19] are used and the received score is divided by the score of an expert CM to produce a ratio as the similarity index. In the "construct-on-scaffold" task the similarity index is calculated by dividing the number of correct answers in blanks by the total number of blanks. The system uses popular schemes described early in the paper, but it does not consider impact of hints used by a student on his/her assessment result.

In [5] the reasonable fallible analyser is developed that compares a student's CM with the expert's one and provides score which is calculated as the similarity index considering fully correct propositions and partly correct propositions in the student's map, as well as valid concepts and number of propositions and concepts in the expert map. Thus, the scoring scheme used in the system considers validity of specific propositions.

An integrated set of assessment tools called HIMATT (Highly Integrated Model Assessment Technology and Tools) [21] calculates descriptive measures for CMs including structural indicators derived from graph theory such as connectedness, ruggedness, number of cycles, etc., as well as two similarity indexes and structural matching, propositional matching, concept matching, and others measures.

One of the most advanced systems is COMPASS [8] offering a range of CM based tasks such as the construction of a map, the evaluation/correction, the extension and the completion of a given map. The analysis of a student's map is performed both quantitatively and qualitatively. Several categories of students' errors are identified and taking into account [8], for example, incomplete relationship, missing relationship, missing concept and its relationships, etc. Weights are defined by a teacher for each concept and proposition in a teacher's CM, as well as for each category of errors. So, the student's score is calculated as the similarity index taking into account concepts and propositions in an expert map and their correctness in a student's map.

CONCLUSIONS AND FUTURE WORK

Regardless of the diversity of schemes for scoring CMs, it is necessary to note that basically they are intended for human-based scoring and developed for the evaluation of tasks, in which learners must create their own CMs. Clearly these tasks belong to the most difficult ones for the development of computer-based CM assessment systems. At the same time, surprisingly but evaluation problems of much more simple fill-in-the-map tasks still remain open despite the fact that fill-in-the-map tasks can be easily embedded in computerized assessment systems and evaluated using an expert map and different quantitative measures. The analysis shows that the existent computer-based CM assessment systems employ rather primitive scoring schemes and, in the best case, they consider only validity of concepts and propositions in students' CM in relation to concepts and propositions in an expert CM. Moreover, several questions have been set aside. For example, it is difficult to evaluate if a student receives valuable information about his/her knowledge level when he/she is presented with information about minimum valence, shortest path, connectedness or similar measures of his/her CM. Even in the most advanced systems such as COMPASS and HIMATT implementing rather rich quantitative and qualitative scoring, assessment of student's knowledge level is not adaptive enough. Such important factors as the level of task difficulty, the number of mistakes made at each level, the frequency with which students use provided help and feedback are not considered at all. In our already implemented KAS the mentioned elements play an important role, so further work will be devoted to the implementation of a scoring scheme based on an abundant student model which includes all identified factors.

REFERENCES

[1] Cañas, A. J. A Summary of Literature Pertaining to the Use of Concept Mapping Techniques and Technologies for Education and Performance Support. Technical report: Pensacola, FL, 2003, 108 p.

[2] Chang, K.E.,. Y. T.Sung, S. F.Chen. Learning Through Computer-Based Concept Mapping with Scaffolding Aid. Journal of Computer Assisted Learning, vol.17, no.1, 2001, pp.21-33.

[3] Chang, K.E., Y.T.Sung, R.B.Chang, S.C.Lin. A New Assessment for Computer-Based Concept Mapping. Educational Technology and Society, vol.8, no.3, 2005, pp. 138-148.

[4] Cimolino, L., J.Kay, A.Miller. Incremental Student Modelling and Reflection by Verified Concept-Mapping. Proceedings of the 11th International Conference on Artificial Intelligence in Education: Learner Modelling for Reflection Workshop, July 20-24, 2003, Sydney, Australia, pp. 219-227.

[5] Conlon, T. Formative Assessment of Classroom Concept Maps: the Reasonable Fallible Analyser. Journal of Interactive Learning Research, vol. 17, no.1, 2006, pp.15-36.

[6] Da Rocha, F.E.L., J.V.Da Costa Junior, E.L.Favero. An Approach to Computer-Aided Learning Assessment. Proceedings of the 3rd International Conference on Concept Mapping, September 22-25, 2008, Tallinn, Estonia and Helsinki, Finland, pp. 170-177.

[7] Goldsmith, T.E., P.J.Johnson, W.H.Acton. Assessing Structural Knowledge. Journal of Educational Psychology, vol.83, no.1, 1991, pp.88-96.

[8] Gouli, E., A.Gogoulou, K.Papanikolaou, M.Grigoriadou. Evaluating Learner's Knowledge Level on Concept Mapping Tasks. Proceedings of the 5th IEEE International Conference on Advanced Learning Technologies, July 5-8, 2005, Kaohsiung, Taiwan, pp. 424-428.

[9] Grundspenkis, J., A.Anohina. Evolution of the Concept Map Based Adaptive Knowledge Assessment System: Implementation and Evaluation Results. Scientific Proceedings of Riga Technical University, RTU Publishing, Riga, 2009 (in print).

[10] Herl, H.E., H.F.O'Neil, G.K.W.K.Chung, R.A.Dennis, J.J.Lee. Feasibility of an On-line Concept Mapping Construction and Scoring System. Paper presented in the Annual Meeting of the American Educational Research Association, March, 1997, Chicago, Il.

[11] Hoeft, R.M., F.G.Jentsch, M.E.Harper, A.W.Evans, C.A.Bowers, E.Salas. TPL-KATS-Concept Map: a Computerized Knowledge Assessment Tool. Computers in Human Behavior, no. 19, 2003, pp. 653–657.

[12] Klein, D.C.D., G.K.W.K.Chung, E.Osmundson, H.E.Herl. Examining the Validity of Knowledge Mapping as a Measure of Elementary Students' Scientific Understanding. CSE Report 557, University of California, Los Angeles, 2002, 45p.

[13] Liu, X. The Validity and Reliability of Concept Mapping as an Alternative Science Assessment when Item Response Theory Is Used for Scoring. Paper presented in Annual Meeting of the American Educational Research Association, New Orleans, LA, April 4-8, 1994, 32 p.

[14] Liu, C.C., P.H.Don, C.M.Tsai. Assessment Based on Linkage Patterns in Concept Maps. Journal of Information Science and Engineering, no. 21, 2005, pp.873-890.

[15] Markham, K.M., J.J.Mintzes, M.G.Jones. The Concept Map as a Research and Evaluation Tool: Further Evidence of Validity. Journal of Research in Science Teaching, vol. 31, 1994, pp.91-101.

[16] McClure, J.R., B.Sonak, H.K.Suen. Concept Map Assessment of Classroom Learning: Reliability, Validity, and Logistical Practicality. Journal of Research in Science Teaching, vol.36, no.4, 1999, pp.475-492.

[17] Mls, K. Implicit Knowledge in Concept Maps and Their Revealing by Spatial Analysis of Hand-drawn Maps. Proceedings of the 2nd International Conference on Concept Mapping, September 5-8, 2006, San José, Costa Rica, 2006, pp.130-133.

[18] Nicoll, G., J.Francisco, M.Nakhleh. A Three-Tier System for Assessing Concept Map Links: a Methodological Study. International Journal of Science Education, vol. 23, no. 8, 2001, pp.863-875.

[19] Novak, J.D., D.B.Gowin. Learning How to Learn. Cambrige: Cambridge University press, 1984.

[20] Pearsall, N. R., J.Skipper, J.Mintzes. Knowledge Restructuring in the Life Sciences: a Longitudinal Study of Conceptual Change in Biology. Science Education, vol. 81, no. 2, 1997, pp. 193-215.

[21] Pirnay-Dummer, P., D.Ifenthaler, J.M.Spector. Highly Integrated Model Assessment Technology and Tools. Proceedings of the IADIS International Conference on Cognition and Exploratory Learning in Digital Age, October 13-15, 2008, Freiburg, Germany, pp. 18-28.

[22] Ruiz-Primo, M.A., S.E.Schultz, R.J.Shavelson. Concept Map-Based Assessment in Science: Two Exploratory Studies. CSE Report 436, University of California, Los Angeles, 1997, 54 p.

[23] Rye, J. A., P.A.Rubba. Scoring Concept Maps: an Expert Map-Based Scheme Weighted for Relationships. School Science and Mathematics, vol.102, no.1, 2002, p.33-44.

[24] Taricani, E.M., R.B.Clariana. A Technique for Automatically Scoring Open-Ended Concept Maps. Educational Technology, Research and Development, vol.54, no.1, 2006, pp.65-82.

[25] Tsai, C.C., S.S.J.Lin, S.M Yuan. Students' Use of Web-based Concept Map Testing and Strategies for Learning. Journal of Computer Assisted Learning, vol.17, no.1, 2001, pp. 72-84.

[26] Vilkelis, M., A.Anohina, R.Lukashenko. Architecture and Working Principles of the Concept Map Based Knowledge Assessment System. Proceedings of the 3rd International Conference on Virtual Learning, October 31- November 2, 2008, Constanta, Romania, pp. 81-90.

[27] Wallace, J.D., J.J.Mintzes. The Concept Map as a Research Tool: Exploring Conceptual Change in Biology. Journal of Research in Science Teaching, vol. 27, no. 10, 1990, pp.1033-1052.

[28] West, D.C., J.K.Park, J.R.Pomeroy, J.Sandoval. Concept Mapping Assessment in Medical Education: a Comparison of Two Scoring Systems. Medical Education, no. 36, 2002, pp. 820-826.

[29] Yao, Q., K.Yang, G.Zhao, R Huang. A Scoring Algorithm Based on Proposition Chains for Concept Mapping. Proceedings of the 2nd International Conference on Concept Mapping, September 5-8, 2006, San José, Costa Rica, 2006, pp.8-15.

[30] Yin, Y., J.Vanides, M.A.Ruiz-Primo, C.C.Ayala, R.Shavelson. A Comparison of Two Construct-a-Concept-Map Science Assessments: Created Linking Phrases and Selected Linking Phrases. CSE Report 624, University of California, Los Angeles, 2004.

ABOUT THE AUTHORS

Lecturer Alla Anohina, Dr.sc.ing., Department of Systems Theory and Design, Faculty of Computer Science and Information Technology, Riga Technical University, Phone: +371 67089595, E-mail: alla.anohina@rtu.lv

Prof. Janis Grundspenkis, Dr.habil.sc.ing., Department of Systems Theory and Design, Faculty of Computer Science and Information Technology, Riga Technical University, Phone: +371 67089581, E-mail: janis.grundspenkis@cs.rtu.lv