Evaluating Students' Concept Maps in the Concept Map Based Intelligent Knowledge Assessment System

Alla Anohina-Naumeca and Janis Grundspenkis

Department of Systems Theory and Design, Riga Technical University, Kalku Str. 1, LV-1658 Riga, Latvia
{alla.anohina-naumeca,janis.grundspenkis}@rtu.lv

Abstract. The paper presents results of the preliminary analysis related to the development of the evaluation mechanism for students' concept maps in the concept map based intelligent knowledge assessment system. Scoring schemes intended for human-based evaluation of concepts maps and embedded in computer-based concept mapping assessment systems are discussed. Overview of the developed system is given and scoring mechanisms implemented in its earlier prototypes are described. Factors affecting students' score in the mentioned system are identified and decisions concerning the development of a new evaluation mechanism are specified.

Keywords: knowledge assessment system, concept maps, scoring mechanism.

1 Introduction

Concept maps (CMs) as an assessment tool represent knowledge in form of a graph which nodes correspond to concepts in a domain, but arcs indicate relationships between concepts. Arcs can be directed or undirected and with or without linking phrases on them. A linking phrase specifies the kind of a relationship between concepts. The main constituent part of a CM is a proposition displaying a relationship between two concepts and corresponding to an elementary unit of knowledge. Usually CMs are represented in a hierarchical fashion [1] and a particular group of hierarchically related concepts is called a segment. Cross-links are relationships between concepts in different segments [2]. Various CM based tasks can be offered to students, however, two main groups of them are: a) "fill-in-the-map" tasks, where the structure of a CM is given to a student and he/she must fill it using the provided set of concepts and/or linking phrases, and b) "construct-a-map" tasks, where a student must decide on the structure of a CM and its content by him/herself.

The Department of Systems Theory and Design of Riga Technical University has been developing a CM based intelligent knowledge assessment system (KAS) since the year 2005. Four prototypes have been already implemented and experimentally evaluated [3]. The current development direction of the mentioned system is the elaboration of an automated scoring mechanism for evaluation of students' CMs. The paper describes evaluation schemes implemented in earlier prototypes of the KAS and presents results of the preliminary analysis concerning a new scoring mechanism.

J. Grundspenkis et al. (Eds.): ADBIS 2009 Workshops, LNCS 5968, pp. 8–15, 2010.

The rest of the paper is structured as follows. Section 2 gives an overview of CM scoring schemes already proposed. Section 3 describes main functionality of the KAS. Section 4 focuses on evaluation of students' CMs in the earlier prototypes of the system, factors affecting students' score and decisions made regarding a new evaluation mechanism. Conclusions are presented at the end of the paper.

2 Related Works

Actually a great number of scoring schemes intended for human-based evaluation have been developed. Our analysis presented in detail in [4] shows that most of them are based on quantitative measures (number of valid propositions, levels of hierarchy, etc.) and only few combine both quantitative and qualitative (categorization of propositions according to the degree of their correctness) approaches. Considering quantitative measures it is difficult to evaluate if a student receives valuable information about his/her knowledge level when he/she is presented with such kind of data. Moreover, the greater part of structural scoring schemes are mainly applicable only for hierarchical CMs because such aspects as levels of hierarchy and cross-links are taken into account. Typically comparing students' CMs with one or more experts' maps a closeness index showing the extent to which the CM of a student matches that of the expert is calculated. The most schemes are developed for the evaluation of "construct-a-map" tasks which belong to the most difficult ones for the development of computer-based CM assessment systems. At the same time evaluation problems of much more simple "fill-in-the-map" tasks still remain open despite the fact that they can be easily embedded in computerized assessment systems and evaluated using an expert map and quantitative measures.

The known computer-based concept mapping assessment systems use rather primitive scoring schemes and in the best case validity of concepts and propositions in students' CM in relation to an expert CM is considered. One of the most advanced systems in this direction is COMPASS [5] offering a range of CM based tasks, performing quantitative and qualitative analysis of a student's map and identifying several categories of students' errors. 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 a similarity index taking into account concepts and propositions in an expert map and their correctness in a student's map.

Unfortunately, a number of important factors (for example, the level of task difficulty, the number of mistakes made at each level, the frequency with which students use provided help and feedback, etc.) are not considered at all in the examined systems. In our already implemented KAS the mentioned elements play an important role and must be taken into account.

3 Overview of the KAS

The KAS has twofold goals: a) to promote students' knowledge self-assessment, and b) to support a teacher in the improvement of learning courses through analysis of results of systematic assessment of students' knowledge. The developed system is used in the following way [6]. A teacher defines stages of knowledge assessment and creates CMs for all of them by specifying relevant concepts and relationships among them in such a way that a CM of each stage is nothing else then an extension of the previous one. During knowledge assessment a student solves a CM based task corresponding to the assessment stage. After a student has submitted his/her solution, the system compares a student's CM with the teacher's one and generates feedback.

At the moment the system provides rich students' support (provided help and feedback) in comparison with other systems [7]. Three kinds of help are supported. Firstly, the system offers 3 "fill-in-the-map" tasks (Task1-insertion of concepts in the structure of a CM containing linking phrases, Task2-insertion of concepts in the structure of a CM without linking phrases, Task3-insertion of concepts and linking phrases in the structure of a CM) and 2 "construct-a-map" tasks (Task4-creation of a CM from the given set of concepts, Task5-creation of a CM from the given sets of concepts and linking phrases). Eight transitions between tasks are implemented allowing a student to find a task most suitable for his/her knowledge level. Four of them increase the degree of task difficulty and other four transitions reduce it. Secondly, in "fill-in-the-map" tasks a student can choose a concept from the given set of concepts and ask the system to insert it into the right place (node) within the structure of a CM. Thirdly, in all previously mentioned tasks a student can choose a concept from the given set of concepts and ask the system to explain it using one of the following types of explanations: definition, short description or example.

Feedback consists of numerical data (maximum score, actual student's score, total time for the task completion, time spent by a student), student's CM marked with labels representing his/her received points for each relationship and possibility to check a proposition. Checking of a proposition is supported at all previously described degrees of task difficulty. A student points out his/her created proposition and the system checks its correctness. Moreover, this feedback in case of incorrectness of a proposition presents explanations of both concepts involved in the proposition as it was described above.

4 Evaluation of Students' Concept Maps

A teacher's created CM serves as a standard against which students' CMs are compared in the KAS. Moreover, a comparison algorithm has been developed which is sensitive to the arrangement and coherence of concepts in students' CMs [8]. The algorithm is capable of recognizing different patterns of a student's solution. Two types of relationships are used in CMs: a) important relationships which show that relationships between the corresponding concepts are considered as important knowledge in a learning course, and b) less important relationships that specify desirable knowledge. For each correctly provided important relationship a student receives 5 points, but for each less important relationship only 2 points are assigned. Each proposition can be evaluated considering relative contribution of its parts: the presence of a relationship in a student's CM - 40%, a correct linking phrase - 30%, a correct direction of the arc - 15%, a correct type - 10%, both concepts related by the relationship are placed in the correct places - 5%.

At the moment the system can recognize more than 36 different patterns of correct and partly correct propositions in students' CMs [8]. Recently the improvement of the mentioned algorithm was made by considering the so called "hidden" relationships in students' CMs which are nothing else than the derivation of relationships presented in a teacher's CM. The hidden relationships are correct too and could appear in students' CMs. They are scored by 1 point.

In general, students' CMs are scored by identifying how many correct relationships a student has defined:

$$P = \sum_{i=1}^{n} p_i * c_i \,. \qquad (1)$$

where P is a student's score after the completion of a task, p_i is the maximum score according to the type of an i-th relationship, c_i is the coefficient that corresponds to the degree of the correctness of an i-th relationship, and n is the number of relationships in the CM structure including hidden relationships.

However, eq.(1) can be used only in case, if a student has completed a task without asking for help. However, taking into account that the system provides several kinds of students' support a correction mechanism must be applied in order to compare: a) results of those students who completed an original task without asking help and those who used help, and b) results of students who performed task at the higher difficulty degree and those who performed the same task at the lower difficulty degree.

Before considering development of the correction mechanism scoring schemes implemented in earlier prototypes of the KAS are described.

4.1 Scoring Schemes Implemented Early

In the year 2006 two versions of the system was implemented. Each of them supported different approach to the changing of the degree of task difficulty: insertion of additional concepts into a CM and offering of different types of tasks.

Thus, in the first approach a task of filling-in a teacher defined CM structure by a given set of concepts was offered to students. During the completion of a task a student could ask to reduce the degree of task difficulty. In this case the system inserted some concepts into the right nodes of the structure of a CM. Two main factors were taking into account developing a correction mechanism [10]: the number of difficulty reduction times and the number of concepts inserted by the system. It was necessary for two reasons. Firstly, concepts inserted by the system facilitated the further solving of a task. Secondly, before the reduction of the degree of task difficulty the system checked a student's solution and only correct or partly correct concepts remained in the CM, while other concepts were removed from it.

The correction coefficient was introduced. Its initial value was 1, but in the case of the reduction of difficulty this coefficient was decreased. The decrease consisted of two parts: a „penalty" for the number of the reduction times of the degree of task difficulty and the proportion of the number of additionally inserted concepts to the total number of concepts within the map. Thus, a student's score was calculated combining eq.(1) with the correction coefficient [10]:

$$P = \left(\sum_{i=1}^{n} p_i * c_i\right) * \left(1 - c_s * s - \sum_{i=1}^{j} \frac{(a + \frac{\Delta * (i-1)}{m}}{m}\right). \tag{2}$$

where c_s is the penalty for each difficulty reduction time, s is the number of difficulty reduction times, a is the penalty for the insertion of the first concept by the system, j is the total number of concepts inserted by the system, m is the total number of concepts in a CM, and Δ is the increase of the penalty for each concept insertion.

In the second approach five tasks described in Section 3 and transitions between them were implemented. So, the previously specified eq.(1) was modified by a coefficient of the degree of difficulty for a given task in the following way [11]:

$$P = \sum_{i=1}^{n} lk_i * p_i * c_i . \tag{3}$$

where lk_i is the coefficient of the degree of difficulty for a given task. The coefficient lk_i was assigned to each relationship, but not to the whole task, because a CM of the current assessment stage could contain relationships defined at the previous stages on different degrees of difficulty. Assignment was made during the completion of a task and depended on the degree of task difficulty [11].

4.2 Factors Affecting Students' Score in the KAS

Improvement of the functionality of the system has lead to the necessity to develop an appropriate evaluation mechanism. However, several important issues appeared. In general, the score of a teacher's CM is calculated taking into account directly created relationships, where each important relationship is weighted by 5 points, but less important – by 2 points. Let's define this score as P_{norm}:

$$P_{norm} = 2 * x + 5 * y . \tag{4}$$

where P_{norm} is the score of a teacher's CM taking into account only directly created relationships, x is the number of less important relationships in a teacher's CM, and y is the number of important relationships in a teacher's CM.

However, as was pointed out before in Section 3, hidden relationships can be revealed in a teacher's CM and they may appear in students' CMs. As a result we can calculate the maximum score (P_{max}) for a teacher's CM:

$$P_{max} = P_{norm} + z . \tag{5}$$

where P_{max} is the maximum score of a teacher's CM taking into account directly created and hidden relationships, and z is the number of hidden relationships.

A student's score (P_s) will be equal with P_{max} in case if a student has related all concepts in the same way as they are related in a teacher's CM and has revealed all hidden relationships. P_s will be equal with P_{norm} in case if a student's CM completely matches a teacher's CM, but he/she has not revealed hidden relationships. Here, the first question arises: How to compare results of students whose score is equal with P_{max} and those, whose score is equal with P_{norm}? Is the latter worse that the former? It is not the case, because he/she has the same structure of a CM as a teacher.

However, the both previously described cases are ideal cases and, as our experience shows, very few students can reach such results. Typically the greater part of student's relationships will be only partly correct taking into account direction of arcs, correctness of linking phrases, etc. Usually, students' score is less than P_{norm}. But in this case we have the next question: How to compare results of students whose score is less than P_{norm} but whose have hidden relationships in their CMs and students whose score is less than P_{norm}, but they do not have extra relationships?

Besides, not only the presence of hidden relationships affects the score of students. Other factors are related to the usage of help: a) students can reduce the degree of task difficulty during the completion of a task; b) students can ask explanations of concepts in all tasks; c) students can check propositions in all tasks, and d) students can ask to insert chosen concepts to right places in "fill-in-the-map" tasks.

Moreover, considering different degrees of task difficulty there is different number of penalties for partly correct propositions. Table 1 shows that working at the 4th degree of task difficulty students always will receive the greater score because there are only 2 penalties, but at the 1st and the 3rd degrees always the less one. This is due to the fact that entry "no" means that for missing parts a student receives full points, for example, at the 5th degree of task difficulty places of all concepts are considered as correct, but in reality places are not important for this task. A problem that arises from the different number of penalties is the following. Working at the 3rd degree of task difficulty a student can place all concepts on correct places, but does not provide linking phrases, and after that to reduce the degree of task difficulty. After moving to the 2nd degree he/she can submit his/her solution and receive the maximum score because linking phrases are not used at this degree.

4.3 Development of a New Scoring Mechanism

Considering the problems mentioned above the following decisions have been made. Firstly, two modes of system's operation must be provided: a) a mode of knowledge self-assessment which purpose is to allow a student to assess its own knowledge level and to learn more about a particular topic in case of incomplete or incorrect knowledge, and b) a mode of knowledge control intended for the determination of students' knowledge level by a teacher.

Secondly, a correction mechanism must be applied in different ways in each mode. During knowledge self-assessment the reduction of a student's score will not be performed in case of usage of such kinds of help as checking of a proposition or explanation of a concept. According to [7] both kinds of support provide not only help in task completion, but also tutoring. Thus, it is not correctly to reduce a student's

Table 1. Number of penalties at different degrees of task difficulty

Difficulty degree	An incorrect place of a concept	An incorrect type of a relationship	An incorrect linking phrase	An incorrect direction of an arc	Total number of penalties
5th	No	Yes	Yes	Yes	3
4th	No	Yes	No	Yes	2
3rd	Yes	Yes	Yes	Yes	4
2nd	Yes	Yes	No	Yes	3
1st	Yes	Yes	Yes	Yes	4

score after usage of such help in the mode of knowledge self-assessment. However, points must be reduced in case of additional insertion of concepts because this kind of help substantially facilitates the task completion particularly if fundamental concepts are inserted. Moreover, it is necessary to avoid situations when a student in such a way inserts all concepts and receives the maximum score. In the mode of knowledge control all kinds of help will contribute to the reduction of a student's score.

Thirdly, it is necessary to define the restriction on the maximum number of propositions which can be checked (only for the mode of knowledge control) by a student and the maximum number of concepts which can be inserted by the system if a student asks for it. Two approaches must be implemented: a) the number of checking allowed must be calculated automatically by the system taking into account the total number of propositions or concepts for a task of the current assessment stage, and b) a teacher's possibility to change the restriction automatically set by the system.

In case of explanation of a concept a student's score must be reduced only in case if a student asks explanation of a certain concept for the first time. After that he/she can receive explanation of the same concept repeatedly without reducing the score.

In the mode of knowledge control penalty for the usage of help must grow each time when a student uses help in order to stimulate his/her to complete a task by him/herself and to use the reduction of difficulty as seldom as possible.

Fourthly, considering a problem related to hidden relationships we decided to provide additional feedback both to a student and to a teacher by showing P_{norm}, P_{max}, P_s, number of hidden relationships in a teacher's CM and revealed by a student.

5 Conclusions and Future Work

An appropriate evaluation mechanism of students' CMs is an important part of any computer-based concept mapping assessment system. Regardless of the fact that a lot of CM scoring schemes have been developed the greater part of them are intended for human-based evaluation of "construct-a-map" tasks. As a result, their feasibility and usefulness in CM based knowledge assessment systems is a discussible question. In turn, evaluation mechanisms implemented in the known assessment systems do not consider such important factors as the level of task difficulty, the number of mistakes made at each level, the frequency of used help, etc. In the authors' developed KAS the following factors affect students' score: the presence of hidden relationships in students' CMs, possibility to reduce the degree of task difficulty, three kinds of help available and different number of penalties at each degree of task difficulty. As a result, the necessity to develop an appropriate evaluation mechanism has arisen. So far, the following decisions have been made: it is necessary a) to provide two modes of system's operation (knowledge self-assessment and knowledge control) and to apply the correction mechanism of students' score in different ways in each mode, b) to implement two approaches (automatic calculation by the system and changing by the teacher) regarding the restriction on the maximum number of propositions which can be checked by a student and the maximum number of concepts which can be inserted by the system, c) to reduce the students' score only in case if a student asks an explanation of a certain concept for the first time, d) to provide growing of penalty for the usage of help in the mode of knowledge control in order to stimulate a student to complete the task by him/herself, and e) to provide additional feedback to a student and to a teacher concerning hidden and directly created relationships. Future works is related to the development of a mathematical model for scoring CMs in the KAS.

References

1. Novak, J.D., Cañas, A.J.: The Theory Underlying Concept Maps and How to Construct and Use Them. Technical report, Florida Institute for Human and Machine Cognition (2008)
2. 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 (2003)
3. Grundspenkis, J., Anohina, A.: Evolution of the Concept Map Based Adaptive Knowledge Assessment System: Implementation and Evaluation Results. In: 49th Scientific Conference at Riga Technical University, pp. 13–24. RTU Publishing, Riga (2009)
4. Anohina, A., Grundspenkis, J.: Scoring Concept Maps: an Overview. In: 10th International Conference on Computer Systems and Technologies, Ruse, Bulgaria, pp. IV.8-1–VI.8-6 (2009)
5. Gouli, E., Gogoulou, A., Papanikolaou, K., Grigoriadou, M.: Evaluating Student's Knowledge Level on Concept Mapping Tasks. In: 5th IEEE International Conference on Advanced Learning Technologies, Kaohsiung, Taiwan, pp. 424–428 (2005)
6. Vilkelis, M., Anohina, A., Lukashenko, R.: Architecture and Working Principles of the Concept Map Based Knowledge Assessment System. In: 3rd International Conference on Virtual Learning, Constanta, Romania, pp. 81–90 (2008)
7. Anohina, A., Grundspenkis, J.: Student's Support in the Concept Map Based Knowledge Assessment System. In: 7th European Conference on e-Learning, Agia Napa, Cyprus, pp. 38–45 (2008)
8. Anohina, A., Vilkelis, M., Lukasenko, R.: Incremental Improvement of the Evaluation Algorithm in the Concept Map Based Knowledge Assessment System. Int. J. Comp., Com.& Cont. 4(1), 6–16 (2009)
9. Grundspenkis, J., Strautmane, M.: Usage of Graph Patterns for Knowledge Assessment Based on Concept Maps. In: 49th Scientific Conference at Riga Technical University, pp. 60–71. RTU Publishing, Riga (2009)
10. Anohina, A., Lavendelis, E., Grundspenkis, J.: Concept Map Based Knowledge Assessment System with Reduction of Task Difficulty. In: 16th International Conference on Information Systems Development, Galway, Ireland, pp. 853–866 (2007)
11. Anohina, A., Pozdnakovs, D., Grundspenkis, J.: Changing the Degree of Task Difficulty in Concept Map Based Assessment System. In: IADIS International Conference e-Learning 2007, Lisbon, Portugal, pp. 443–450 (2007)