Chen, W. et al. (Eds.) (2016). Proceedings of the 24th International Conference on Computers in Education. India: Asia-Pacific Society for Computers in Education

Comparison of Concept Map Evaluation between Kit-Build Method and Handmade Method

Warunya WUNNASRI*, Jaruwat PAILAI, Yusuke HAYASHI & Tsukasa HIRASHIMA
Graduate school of Engineering, Hiroshima University, Japan
* warunya@lel.hiroshima-u.ac.jp

Abstract: This paper describes the reliability of our automatic concept map evaluation framework that is called Kit-Build concept map. This framework is developed based on concept map that is used for organizing and representing knowledge. It is enhanced to support an education. However, the automatic concept map evaluation has not investigated about the reliability. To confirm the reliability of Kit-Build concept map, we compare Kit-Build concept map with two handmade concept map evaluation methods that contain the structural scoring and the propositional scoring. These handmade methods can evaluate concept map flexibly because the human can understand the meaning of each proposition in concept map even the words of proposition do not appear in a learning material. So the handmade methods are claimed that they have the reliability enough for evaluating concept map. To reduce time cost and human workload, researchers propose the automatic concept map evaluation that becomes an important and useful for a classroom situation. It can evaluate a lot of concept maps in a short time, but it still needs to be examined the reliability. We designed preliminary experiment in two learning situations that are teaching and reading situations and compared the correlation between the handmade methods and Kit-Build concept map. Even though these are preliminary results, they suggest that using Kit-Build concept map in teaching situation gets acceptable reliability when it is compared the correlation with two handmade concept map evaluation methods.

Keywords: Concept Map Evaluation, Kit-Build Concept Map

1. The Concept Map Evaluation Method

Joseph D. Novak developed concept map as a graphical tool for representing and organizing knowledge, and it is also used to evaluate learners understanding in classes. So a lot of handmade concept map evaluation methods are proposed and get the reliability widely because their procedures require the human for evaluating. Nevertheless, these methods take cost such as time cost and human workload too much, so it is not convenient to use in classes that have a lot of learners. Thus the automatic concept map evaluations are proposed, but they still have to be investigated the reliability.

1.1 The Handmade Concept Map Evaluation Method

From our investigation, the handmade concept map evaluation methods are categorized by using the precedence of each method. One of a typical method is the Novak and Gowin structural method (Novak & Gowin, 1984), which is grouped in the structural scoring. They give high scores the precedence on a level of the hierarchy and the number of crosslink in a concept map while the valid proposition can get only one score per proposition. The significant meaning of proposition may be neglected. That is different from the McClure and Bell relational method (McClure et.al, 1999). This method is grouped in the propositional scoring because it pays attention to the meaning of proposition precedence. The procedure investigates the suitability of meaning of each proposition. If the linking word is appropriate with concepts clearly, that proposition will get three scores as a perfect score. The score will be depreciated depending on the meaning of linking word. It is reasonable for evaluating understanding from concept map. But these handmade methods have to use an expert for evaluating and require a long time for scoring each concept map.

1.2 The Automatic Concept Map Evaluation Method

Because of the time consuming and workload that has to pay for the handmade evaluation, the automatic concept map evaluation methods are proposed. Most of them use the criteria map as the target of learning. They compare the learner map with criteria map to evaluate learners' understanding. In this study, we call the automatic comparison concept map evaluation method. This comparison inherits the property from the human method that is the structural scoring and propositional meaning scoring. If learner maps are same as criteria map, it shows that learners can understand in instructor's objective well, which includes the understanding of structure and meaning of the proposition.

Our framework, Kit-Build concept map (Hirashima et al., 2015) is an automatic concept map assessment that uses the exact matching in propositional level for evaluating concept map. It has been already employed in classrooms practically and confirmed that the framework and results of the diagnosis were useful to support teachers in science learning in elementary school. That proves that it is suitable for using in teaching situation that instructor gives the direction following instructor's interpretation. However, we have not examined the quality of the propositional exact matching evaluation. And we have not studied our framework in reading the situation that learners have to interpret material by themselves. So we produce the experiment to investigate the reliability of Kit-Build concept map by comparing well-known handmade evaluation methods. For using Kit-Build concept map, the instructor has to prepare the criteria map, which is called the goal map in our framework. It is constructed as the informal concept map because it should follow the instructor's objective that requires learners to understand that is not the universe context. After that, the goal map is extracted to the kit that contains a list of concepts and relationships. This kit that is provided to learners can help learners to reduce their cognitive load more than the traditional concept map, which they must create all components by themselves. After that, learners are requested to reconstruct concept map by using the kit, and it is called the learner map. The framework will check leaner maps by exact matching on each learner's proposition with goal map's proposition and generates a similarity score. The instructor can investigate learners' misunderstanding individually and can find the overview of all learners by overlaying concept map as the group map and the group-goal difference map immediately. After result analyzing, the instructor can adjust the goal map or teach learners about leaky content again.

2. Research Methodology

To confirm the reliability, we produce the preliminary experiment to compare the correlation between the handmade concept map evaluation method and Kit-Build concept map. For the handmade evaluation method, we chose the Novak and Bell structural concept map evaluation and the McClure relational propositional method that they are the typical traditional method.

In this preliminary experiment, ten university students were requested to read the article that described "Introduction of concept map." After that, they had to construct concept maps following their reading interpretation by using 21 provided concepts on CmapTools application. It means they must create linking word by themselves. Two handmade evaluation methods evaluated these concept maps, and the raw scores of each method are normalized by using their perfect score. Then, they had to use Kit-Build concept map to reconstruct concept map by using kit. The kit contained 21 concepts that are same as provided concepts in CmapTools and additional 22 relationships. These concept maps were evaluated by our automatic evaluation method. The score is represented as the similarity score when learner map was compared with the goal map. After reading situation, the instructor taught the participants about the same article following instructor's interpretation. And the participants were requested to construct the concept map as same as steps in the reading situation. They had to create linking words by themselves and Kit-Build concept map.

3. Results of Preliminary Experiment and Conclusions

From the preliminary experiment procedure, the average scores of each evaluation method in reading and teaching situation are shown in Table 1. All scores of the reading situation are lower than teaching

situation because participants used their interpretation to create concept maps that are the different way from the instructor's interpretation. And the score of the propositional scoring is higher than the structural scoring because the evaluator tried to understand the meaning of each proposition and tried to give the score to participants as much as possible. That is different from the structural scoring that forces evaluator to give the precedence to the structure of concept maps too much. It can make some leaky essential meaning. So the McClure and Bell relational method can show learners understanding more meaningful than the Novak and Gowin structural method.

Table 1: The result of experiment (n=10)

	Structural Scoring (Novak and Gowin's)	Propositional scoring (McClure and Bell's)	Automatic comparison (Kit-Build concept map)
Reading	24.56	35.09	30.30
Teaching	28.99	53.41	50.00

While the score of Kit-Build concept map is in the middle area, it is lower than the McClure and Bell propositional scoring but is higher than the Novak and Gowin structural scoring. Even though Kit-Build concept map uses the exact matching to score each proposition of learner maps, the scores are acceptable when are compared with the handmade concept map evaluations, which use very flexible matching for scoring the maps. Table 2 shows the correlation between the score of handmade methods and Kit-Build concept map. The p-values show we cannot discuss the correlation between the handmade methods with Kit-Build concept map in reading situations. Because, when the participants read the article, they interpreted the article by themselves and it is possible to be various ways. While the result from teaching situation has a marginal medium correlation between both handmade evaluation method and Kit-Build concept map, it shows the lecture from instructors can make an agreement on the article by teaching and guides the learners to understand in the same direction with the instructor. From the assumption that the handmade evaluation methods are reliable, we conclude Kit-Build concept map has reliability enough for evaluating concept map in teaching situation because it is not much different when it is compared with reliable handmade concept map evaluation methods.

Table 2. The correlation between handmade evaluation method and Kit-Build concept map

	Kit-Build in reading situation	Kit-Build in teaching situation
Novak and Gowin's structural scoring method	0.1406 (p-value=0.6984)	0.6209 (p-value=0.0553)
McClure and Bell's relational method	0.2702 (p-value=0.4503)	0.5520 (p-value=0.0980)

For the future work, we plan to make the full experiment and compare the stability of handmade evaluation and Kit-Build framework. Because different evaluators may score concept map by various interpretation individually, the score of handmade evaluation will depend on evaluators' interpretation that affects the stability of the assessment. While Kit-Build concept map uses the exact matching for comparing between goal map and learner map, it will return the same result in anytime. From this assumption, we try to confirm Kit-Build concept map has ability enough for evaluating concept maps.

Acknowledgements

This work was supported by JSPS KAKENHI Grant Number 15H02931.

References

Hirashima, T., Yamasaki, K., Fukuda, H., & Funaoi, H. (2015). Framework of kit-build concept map for automatic diagnosis and its preliminary use. *Research and Practice in Technology Enhanced Learning*, 10(1), 1-21.

McClure, J.R., Sonak, B. & Suen, H.K. (1999). Concept map assessment of classroom learning: Reliability, validity, and logistical practicality. *Journal of Research in Science Teaching*, 36(4), 475–492.

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