Concept Maps: Theory, Methodology, Technology
Proc. of the Second Int. Conference on Concept Mapping
San José, Costa Rica, 2006

MEASUREMENT AND EVALUATION METHOD FOR A CONCEPT MAPPING TEST BY DRAWING ORDERING RELATIONS AMONG CONCEPTS

Makoto Takeya, Nobuaki Yasugi, Yoshio Funabashi & Keizo Nagaoka, Japan
Takushoku University, Mitsubishi EIN, Fujitsu SSL & Waseda University, Japan
Email: takeya@cs.takushoku-u.ac.jp

Abstract. The traditional paper test is able to measure mainly individual student's understanding level of individual bits of knowledge. However, it is difficult to measure the internal connection among bits of knowledge. Previously, Takeya, *et al.* (2004) had presented a new testing method, called a concept mapping test for a formative evaluation tool at the 1st CMC. This is a test by using concept maps based on the prerequisite relations among concepts. However, future subjects lie in an evaluation method for a concept mapping test based on ordering relations with transitivity law, such as casual relations, inclusion relations, and so on. This paper presents new measurement and evaluation method for these kinds of structural knowledge. Secondly, giving a few examples, this paper discusses scoring method based on qualitative degree of sequencing. Thirdly, these characteristics and validities of the measurement and evaluation method are discussed.

1 Introduction

Lectures and texts are arranged in linear or sequential order. Each learning unit is presented in order. That is, they naturally move from one idea to the next, and so forth, without ever systematically detailing the structural relationships among these ideas. The teacher is concerned with assessing and promoting the acquisition of knowledge by individual students. Attention has recently been drawn to what has become known as 'structural knowledge' or knowledge of interrelationships among ideas in their knowledge domain. The authors are becoming aware of the need to establish the internal connectedness of ideas and concepts to be learned. It is difficult to evaluate these internal relationships among ideas by using traditional paper tests, because these tests mainly measure the understanding level of individual bits of knowledge obtained by individual students.

2 Measurement and evaluation by a concept mapping test based on ordering relations

It is very important to check whether each relationship, *i.e.* each arrow in the student's map exists or not compared with the teacher's map. Especially, we have to investigate whether misunderstanding of relationships occur or not under the influence of existence or non-existence of individual arrows. Define a concept map by a digraph (directed graph) $G=(V,E)$, where V is a set of vertexes $V=\{v_1,v_2,\ldots,v_n\}$, and E is a set of arrows $E=\{e_1,e_2,\ldots,e_m\}$. The remarkable feature of the concept map G with respect of casual relations, inclusion relations, ordering relations, and so on is that the map is transitive. Here, G is transitive if and only if there exists an arrow (v_i,v_k) whenever both arrows (v_i,v_j) and (v_j,v_k) exists. That is, it is satisfied that whenever $v_i \to v_j$, and $v_j \to v_k$, then $v_i \to v_k$. Previously, Takeya (1999) and Takeya, et al. (2004) had presented the similarity index concentrating on relations of arrows. Contrary to this, a new similarity index should be discussed from a view point of ordering with transitivity law. Next, an example of a concept mapping test is shown before the new similarity index will be introduced in 2.2.

2.1 A concept mapping test on casual relationships

A concept map has been utilized in lectures on Environmental Science at Takushoku University in Japan. After lecturing and showing a video on the "Crisis of the Living Environment at the Foot of Mt. Fuji", a performance test is given by use of a concept map. Test contents are shown as follows:
Included are the following nineteen elements. Please draw a concept map, where an arrow "$a \to b$" means a direct relation between cause a and effect b. Please show your map in the form of a hierarchical structure. Here, the element on the first level is only the element (1).

As shown in Figure 1, their contents include the following:

(1) Crisis of the living environment at the foot of Mt. Fuji (the target)
(2) Industrial use of underground water
(3) Water pollution in rivers
(4) Appearance of artificial valleys
(5) Agricultural damage
(6) Atmospheric contamination
(7) Change of underground water to salt water
(8) Deforestation
(9) Sprinkling of agricultural chemicals
(10) Factory construction
(11) Increase of diseased trees in forest
(12) Dumping of factory wastes into rivers
(13) Gas emissions from automobiles
(14) Decrease of underground water
(15) Decrease of water retention power of the ground
(16) Golf course development
(17) Occurrence of avalanche accidents
(18) Subaru-line road construction
(19) Dumping of factory wastes into rivers

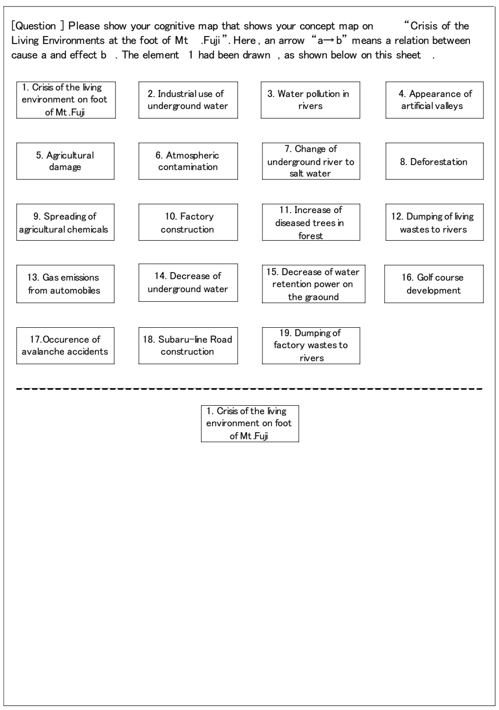

Figure 1. An example of the concept mapping test sheet.

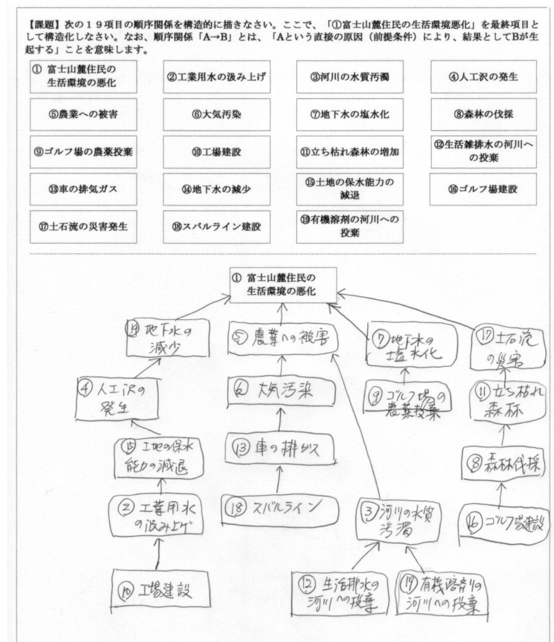

Figure 2 An original sheet of a concept mapping test drawn by one of students. (in Japanese)

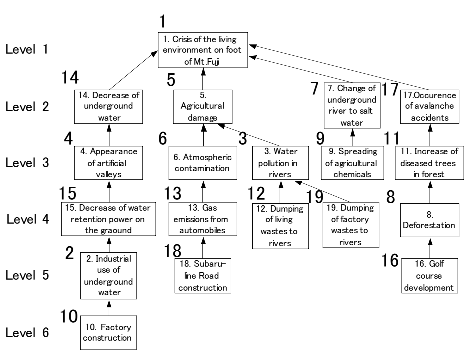

Figure 3 The rearranged English version of the concept map shown in Figure 2.

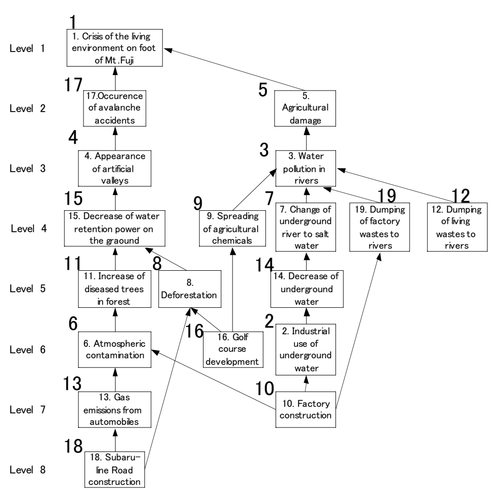

Figure 4 The concept map drawn by a professor.

Figure 1 shows a Japanese original of a concept mapping test drawn by one of students. Figure 2 represents rearranged English version of the concept map shown in Figure 1. Corresponding to Figure 3, Figure 4 shows the concept map drawn by the professor. According to the calculation of the performance score described in 2.2, the professor can understand the performance level for each student. In this case, the student receives a score of 42, described in 2.2.

2.2 *Measurement and evaluation method*

In the case where the digraph satisfies the transitive law, such as the map on casual relationships, and inclusion relationships, the teacher has to give attention to ordering of vertexes. Call a transitive digraph except a totally ordered digraph as a partially ordered digraph. So, represent both an antecedent set of the vertex v_j and a reachable set of v_j on G_X, by $V\left\langle \to v_j \right\rangle^X$ and $V\left\langle v_j \to \right\rangle^X$, respectively. Then, similarity index $S'(G_T, G_X)$ among the teacher's map G_T and the student's map G_X is defined as follows:

$$S'(G_T, G_X) \equiv \frac{\sum_{j=1}^{n} \left| \left(V\left\langle \to v_j \right\rangle^T \cap V\left\langle \to v_j \right\rangle^X \right) \cup \left(V\left\langle v_j \to \right\rangle^T \cap V\left\langle v_j \to \right\rangle^X \right) \right|}{\sum_{j=1}^{n} \left| \left(V\left\langle \to v_j \right\rangle^T \cup V\left\langle \to v_j \right\rangle^X \right) \cup \left(V\left\langle v_j \to \right\rangle^T \cup V\left\langle v_j \to \right\rangle^X \right) \right|}. \qquad (1)$$

Naturally, the following inequality is obtained.

$$0 \le S'(G_T, G_X) \le 1 \qquad (2)$$

Considering the maps in Figure 1 as partially ordered digraph, let's show examples of calculation process. For example, pay attention to the vertex 3.

$V\left\langle \to 3 \right\rangle^T = \{2,7,9,10,12,14,16,19\}$, $V\left\langle \to 3 \right\rangle^X = \{12,19\}$,

$V\left\langle 3 \to \right\rangle^T = \{1,5\}$, $V\left\langle 3 \to \right\rangle^X = \{1,5\}$, and

$\left| \left(V\left\langle \to 3 \right\rangle^T \cap V\left\langle \to 3 \right\rangle^X \right) \cup \left(V\left\langle 3 \to \right\rangle^T \cap V\left\langle 3 \to \right\rangle^X \right) \right| = 2 + 2 = 4$

Next, for the vertex 8, $V\left\langle \to 8 \right\rangle^T = \{16,18\}$, $V\left\langle \to 8 \right\rangle^X = \{16\}$,

$V\left\langle 8 \to \right\rangle^T = \{1,4,15,17\}$, $V\left\langle 8 \to \right\rangle^X = \{1,11,17\}$, and

$\left| \left(V\left\langle \to 8 \right\rangle^T \cap V\left\langle \to 8 \right\rangle^X \right) \cup \left(V\left\langle 8 \to \right\rangle^T \cap V\left\langle 8 \to \right\rangle^X \right) \right| = 3$.

In the same way as the vertex 3 and 8, these values of all the other vertexes involved in G_T and G_X are calculated.

2.3 *Characteristics and validity of the similarity index*

Here, to analyse similarity index, $S'(G_T, G_X)$ is linearly transformed as follows:

$$\psi(G_T, G_X) \equiv 2S'(G_T, G_X) - 1. \qquad (3)$$

That is, $-1 \le \psi(G_T, G_X) \le 1$. (4)

Let's refer to this index as the *Takeya's* ψ. According to *Eq.*(3), it is considered that this index has very interesting characteristics. Ranking data can be interpreted as vertexes on a linear digraph such as G_L shown in Fug.5. Strictly speaking, note that the linear digraph is a totally ordered digraph which holds the transitive law. Whenever both G_T and G_X are linear graphs, then the *Takeya's* ψ coefficient is equivalent to both the *Goodman-Kruskal* γ coefficient, *Somers' d* coefficient and *Kendoll*'s τ coefficient well known as measures of association in ordinal data.

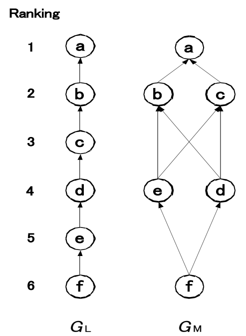

Figure 5 Graphical structure of totally ranking data .

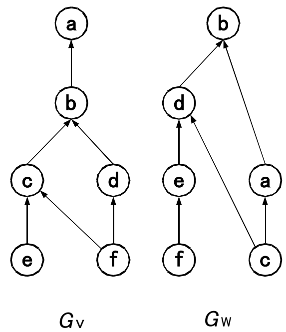

Figure 6 Graphical structure of partially ranking data.

That is, whenever both G_T and G_X are linear graphs, then

$$\psi(G_T, G_X) = \gamma(G_T, G_X) = d(G_T, G_X) = \tau(G_T, G_X)\ . \tag{5}$$

Moreover, in the case where γ coefficient is defined in the expanded range of non-linear data structure as shown in Figure 6, the following relation is obtained.

$$\gamma(G_T, G_X) \geq \psi(G_T, G_X)\ . \tag{6}$$

The proof of the *Eq*.(5) and (6) are omitted by the space limitation. If examined in detail distributions of ψ and γ of several practical concept mapping tests, the values of *Takeya's* ψ coefficient cover the range from -0.5 to 1, but the most of values of *Goodman-Kruskal* γ coefficient converge on the range from 0.8 to 1. Also, in Figure 6, $\gamma(G_V, G_W) = 1.00$, but $\psi(G_V, G_W) = .40$. The ordered pairs $(d, a), (e, a), (f, a), (e, c)$ and (f, c) exist on G_V, but never exist on G_W. On the other hand, the ordered pairs $(c, d), (e, d)$ and (f, e) exist on G_W, but there never exist on G_V. The γ coefficient is taken no account of the above differences. On the other hand, those differences are built in to the ψ coefficient. Therefore, it is considered that the ψ coefficient is the rank correlation generalized to partially ordered data structure. Here, this theoretical description is omitted. As a result, the ψ coefficient is applicable not only to linear hierarchical structure, but also to non-linear one. That is, the *Takeya's* ψ coefficient is generalization of traditional several rank coefficients.

3 Summary

This paper presented measurement and evaluation for concept mapping tests based on casual relationships. Secondly, the measurement method for structural knowledge was shown by using both the models of a concept map and the actual mapping tests. Thirdly, this paper discussed scoring methods according to qualitative degree of sequencing. Lastly, the validity was confirmed. Especially, it was derived that the scoring method for partially ordered map is based on the *Takeya's* ψ coefficient which is generalized from the traditional *Goodman-Kruskal* γ. One of the problems to be solved in near future is development on a reformation algorithm for individual students' misunderstanding.

4 References

Takeya, M. (1999). Structure analysis methods for instruction: Theory and practice of instructional architecture, design and evaluation, Takushoku University Press, Seizansya.

Takeya, M., Sasaki, H., Nagaoka, K., & Yonezawa, N. (2004). A performance scoring method based on quantitative comparison of concept maps by a teacher and students, In A. J. Cañas, J. D. Novak & F. M. González (Eds.), *Concept Maps: Theory, Methodology, Technology. Proceedings of the First International Conference on Concept Mapping* (Vol. 2, pp. 343-346). Pamplona, Spain: Universidad Pública de Navarra.