P1.37

A METHOD OF FORMATIVE EVALUATION NECESSARY FOR EFFICIENT LEARNING

Kohki ITOH* Hisashi SHIO**
* Hokkaido University of Information Science, Ebetsu, Japan
** Institute of Atmospheric Electricity of Hokkaido, Sapporo, Japan

1. INTRODUCTION

To meet increasing demands and expectations, a new system of weather forecasting was introduced in Japan as a national standard in 1993. The occupation of weather forecaster also came to be seen as a desirable job, with many people attempting to pass the rigorous test and attain the qualifications necessary to pursue this career. The pass rate has typically been around 9% of the total number of applicants. Clearly, an efficient learning program is needed. It is also important that the learner can receive progress feedback through formative evaluation.

In this paper, we propose a procedure for formative evaluation. Our method focuses on identifying the order of achievement of learning tasks, based on correct/incorrect responses to test items. Examining the level of achievement after completing a given task implies that the learner's current level of functional ability can be measured.

2. FORMATIVE EVALUATION

Meteorology covers a number of wide and complex scientific fields. Students need to master various disciplines, from basic to applied concepts, as forecasting weather involves making decisions based on a myriad of information. An insufficient or incorrect understanding of a certain learning item will result in incorrect forecasting.

Thus, with the possible exception of elementary concepts, it is very difficult to gain a correct and systematic understanding of meteorology through simple rote memorization. It can also be argued that an insufficient understanding of a fundamental item or concept hinders the complete learning process. It is therefore essential that teachers and learners themselves are able to identify unmastered concepts.

* Kohki ITOH, Hokkaido Univ. of Information Science, Dept. of Information Science, Hokkaido, 069–8585 Japan; e-mail: itoh@do-johodai.ac.jp

** Hisashi SHIO, Institute of Atmospheric Electricity of Hokkaido, 004–0866 Japan.

As a method for evaluating the learner's present level of achievement, written examinations are convenient and enjoy wide popularity among the examiners. They are especially useful in evaluating the current status of a large number of learners at a given time. Test items include appropriate learning tasks, with the learner's current level of achievement being estimated from the test results.

B.S. Bloom grouped learning evaluation into three systematic classes; diagnostic evaluation to diagnose readiness, summative evaluation to appraise the whole course upon completion, and formative evaluation, aiming at the construction of a feedback system for the learner and teacher.

Clearly, learning evaluation is useful for not only the learner but also the teacher, as it allows the latter a ready method to evaluate the appropriateness of the educational material and method of instruction. For example, if a significant number of learners show a similar tendency when tested on a specific concept, it is probable that the method of guidance itself may be a problem.

When learners are unable to complete a certain task, they can be divided into specific two groups; one group recognizing the cause (of being unable to complete the task), and the other group not doing so. In the case of the latter group, it is difficult to determine the cause and the best thing that the teacher can do in such circumstances is made an educated guess. This is clearly far from optimal.

An obvious method to ascertain the cause would be to counsel every learner. However, constraints brought about by time and available interviewers tend to make this too difficult to put into practice. While statistical tools may be helpful in a general evaluation, providing such data as averages and variances, it is only helpful in understanding the general tendency of a group of learners, and not in clarifying each individual problem.

Therefore, what is needed is a method of evaluation using not only the ratio of correct to incorrect answers, but a way to view the pattern of correct/incorrect responses. Teachers generally estimate the order of achievement of test items based on the pattern, and compare the order of achievement within a logical structure of learning tasks, thereby judging the level of ideas formed by the learners. Typical methods utilizing this procedure include Takeya's Item Relational

Structure (IRS)(Takeya, 1980) and Itoh's Fuzzy Item Relational Structure (FIRS)(Itoh, 1994).

3. IRS & FIRS

It can be supposed that a learner makes incorrect answers due to an incorrect understanding of the associated concept. If this incorrectly formed concept sounds logical to the learner, it is accepted and retained by the learner. It is important for the teacher to be able to identify this misunderstood concept, as doing so enables a remedial program or support structure to be appropriately established.

To demonstrate, at first we consider any two test items, a fundamental test item t_i and an applied test item t_j. It is assumed that achievement or understanding of item t_i is necessary to proceed to item t_j. A contingency table is shown in Figure 1, where n is the number of learners, divided into categories a, b, c and d. As category c approaches zero, the assumption that $t_i \rightarrow t_j$ is satisfied.

		t_j 1	0	Total
t_i	1	a	b	a+b
	0	c	d	c+d
	Total	a+c	b+d	n

1: Correct Answer
0: Incorrect Answer

Figure 1: A CONTINGENCY TABLE

To aid in evaluation, we have developed FIRS analysis (Fuzzy Item Relational Structure), a method whereby not only the learning traits of respective learners and groups can be analyzed, but common traits of the whole class as well. Data obtained from test scores is expressed as having a value of [0, 1].

4. RESPONSE AND ACCURACY RATE

The above mentioned analysis method simply classifies incorrect answers and non-responses into an identical category. Strictly speaking, these should probably be classified into different categories. It is often the case that the learner is given sufficient time to answer in a formative test. This implies that there is clearly some difference in the quality of understanding between the two groups.

It is possible to assume some causes of the non-response or incorrect answer. As a representative cause, failure of the learning and acquisition of incorrect concepts are considered. If a concept is formed incorrectly, or misunderstood, it can clearly be expected that there will be some incorrect responses based upon the subject's understanding of the concept.

Figure 2 shows the relation between response rate and accuracy rate. Response rate is the number of responses divided by the number of test items, while accuracy rate is the number of correct answers divided by the number of responses. A cluster represents a group of learners demonstrating similar features.

FIRS analysis typically divides the learner into a cluster based on the reaction pattern to a test item. Additionally, in this paper, a further procedure is added whereby the incorrect answer is distinguished from the non-response.

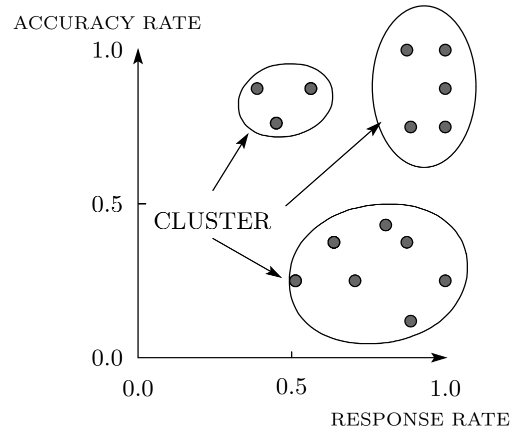

Figure 2: RESPONSE AND ACCURACY RATE

5. CONCLUSION

The teacher guesses a student's cognitive structure from the pattern of the error. We propose that IRS or FIRS analysis should be carried out in every cluster, and suitable learner models should be designed appropriate to the logical structure of the learning tasks. IRS or FIRS are an evaluative method which allows the learner to be quantitatively analyzed in terms of level of understanding. Just as this method has shown to be valuable in assisting both teacher and student to develop a way to evaluate understanding in meteorological education programs, it is envisaged that this will be most beneficial in on-line education.

REFERENCES

Bloom,B.S., 1971: Handbook on formative and summative evaluation of student learning, McGraw-Hill.

Itoh,K. and Ohuchi,A., 1994: Fuzzy Item Relational Structure Analysis for Evaluation of Students by Performance Scores, *Information Processing Society of Japan*, Vol.35, No.5, 945–954.

Takeya,M., 1980: Construction and Utilization of Item Relational Structure Graphs for Use in Test Analysis, *Journal of Educational Technology*, Vol.5, No.3, 93–103.