UNITECH '10 GABROVO

INTERNATIONAL SCIENTIFIC CONFERENCE
19 – 20 November 2010, GABROVO

IMPROVING E-LEARNING WITH DIAGNOSTIC ASSESSMENT

Lilyana Nacheva-Skopalik
Technical University of Gabrovo

Abstract

The modern e-learning platforms still need improvements in order to become effective learning systems that provide high quality learning. One of the solutions in this direction is to make diagnostics of students' learning status and to use the diagnostic results for supportive feedback to students. The diagnostic test provides diagnostics at two levels – individual student and a student group. It is part of cumulative test system for formative assessment that is used for optimal automated control of the leaning process. The test system is also effective assistive tool for teachers.

Keywords: diagnostic assessment, formative assessment, automated feedback, logical control, optimal learning, cybernetic learning model

INTRODUCTION

E-learning is has become integral part of the modern education. There are a lot of e-learning platforms; however, most of them still need further development to become effective learning environments providing quality education. There are different ways to solve this problem. One of the possible solutions is to apply effective formative and diagnostic assessment. This is a vital component to achieve quality learning providing individual learning support and motivation to students and it is estimated as effective assistive tool for teachers in their pedagogical load. Formative assessment and diagnostics are not a common practice for the universities and there is practical need to incorporate it as an integral part of the curricula. Assessment affects the student approach to learning; it is a complicated and contradictory component of learning where more research is needed for innovative and effective solutions.

CYBERNETIC LEARNING MODEL

Our research suggests that in order to develop an optimal and effective learning system it is necessary to include the main features of cybernetic control principles in the system. The learning process in its variety, regardless of the mode of study includes the main characteristics that define it as a cybernetic process. It is possible to use the feedback channel in the cybernetic learning model for optimal control of the learning.

Developed cybernetic learning model considers the hierarchic structure of education and includes control of the learning process at two levels where the objects of control are respectively the individual student and the student group.

The cybernetic approach for optimisation of the learning process requires solving of various tasks, including identification of the object of control. In practice this means to define the student learning characteristics. That is a version of a student model for the individual student and for the student group. A combined approach is developed to define the priorities and weighting coefficients of educational characteristics using an objective evaluation of subjective expert opinions by rank correlation methods and elements of subjective statistics. The chosen characteristics are: level of knowledge, level of skill, total time for learning, intermediate time periods for mastering i^{th} topic, type of the incorrect/non answered questions. They are estimated as sufficient for formative and diagnostic assessment (FDA) and control of the learning process, using them separately or in combination.

Our research develops strategy for optimal control of learning that includes analysis of all learning characteristics and their comparison with the accepted norms. This is in practice diagnostics of the student learning status. The feedback to students to support their learning

and the control decisions depend on diagnostic results.

The specified learning characteristics are obtained by the developed specific cumulative test system. It realizes also the feedback channels for control of the learning process within the cybernetic learning model. It uses logical control algorithm and provides automatically most of the data analysis, diagnostic messages and control decisions. This is considered to be intelligent feature of the cybernetic learning model [3].

CUMULATIVE TEST SYSTEM FOR FORMATIVE AND DIAGNOSTIC ASSESSMENT

The developed cumulative test system for formative and diagnostic assessment (FDA) consists of: initial test (base-line test); self-control tests; tests for knowledge evaluation (TKE); tests for knowledge evaluation and diagnostics (TKED); final test (written exam). As mentioned above, the test system conforms to the accepted approach of programmed learning. Initially, it was developed for technical subjects; however the experience shows that it is suitable for various disciplines and modes of study [4]. The paper presents only the TKED.

TEST FOR DIAGNOSTIC ASSESSMENT

In education "diagnostics" is: methods and means to find out the status of the learner and/or the process of learning. The suggested variant for diagnostics does not present a full diagnostics of the learning and the learner; however it gives information that is sufficient for the purpose of formative assessment and control of the learning process. Because formative assessment is not connected with a crucial final grade for the course any advice for improving the learning is useful.

Diagnostic tests have to define the achieved level of knowledge for particular teaching material, trace the student's learning condition as well as to detect arising problems and the possible reasons for them. The diagnostic results are basis for detailed supportive feedback to students and for making relevant control decisions to stimulate and motivate them to improve their learning. The development of diagnostic tests is a complex task. Further more, it is even more difficult in open and distance learning and e-learning where there is a lack of face-to-face teacher-learner contact and in most cases the contact is indirect, through the available communication tools. That is why making a full and qualitative analysis and diagnostics of the student learning condition is an extremely difficult task. The developed version is estimated as sufficient for the purpose of optimal learning control. The TKED has the same structure as the test for knowledge evaluation and there are no additional difficulties for the students when they conduct the test.

The level of knowledge is analyzed in two versions: static and dynamic.

Static analysis of the level of knowledge for individual student

The static analysis is in practice classifying the test score according to the current grading scale. All students receive information for their score (grade) for each test.

Static analysis of the level of knowledge for student group

Specific adapted version for analysis of histogram is developed. It conforms to the specific features of the learning process. Our experience shows that most of the histogram shapes can appear during the learning process.

Bell-shape histogram shows normal learning process with possibilities for improvement. The control impacts to the group here could be: encouragement, stimulation and other positive impacts. For skewed right (positively skewed) histogram the mean is moved to the left and the shape is asymmetric. This is a case when the student group is weak and it is necessary to actively support it. For skewed left (negatively skewed) histogram the mean is moved to the right. For education this is a signal that the group is relatively good. It is not necessary a special influence for traditional normal teaching courses. Only for cases where extremely good results are aimed (for instance for competitors) some stimulating impacts would be necessary. Plateau-shape histogram and histogram with two peaks (bimodal) generally appear when

the data from a few processes with different mean are mixed. For such cases it is recommended to separate the data based on the individual processes and to explore the reasons for the dispersion. For education this situation could be a signal that the student group is not well formed or a few groups with different capabilities are mixed. Administrative impacts (discussions, meetings, deeper analysis of the student's status etc.), are necessary to clarify the situation. Very often a reason for histogram with outliers (isolated peak) is that data from other process are included. The possible reason for such a situation in education may be using different evaluation criteria (for instance two teachers). When an automated test is used, though, this reason is ignored, because the subjective factor is excluded from the evaluation process. If this histogram type appears the necessary measures are similar to these for bimodal histogram [2].

Dynamic analysis of the level of knowledge

The dynamic analysis of the parameter level of knowledge is in practice analysis of the student performance trend. The same approach is used for analysis of the performance trend for the individual student and for the whole student group.

Using control charts for tracing the "history" of the process is a well known statistical approach applied in quality control. The control chart, in the case of the learning process, is the student performance trend. The developed analysis of the performance trend is adapted to the specific features of the learning process [2].

Two characteristics are used in analysis of the student performance trend: the average value for the test results and the difference between the test results. For the individual student the average value k_{av} is calculated using the accepted approach of cumulative average (moving average). The score from the current test and the two previous tests are used. The same three values are used to analyze the student performance trend. Analysis of the average value includes:

- Computing the average value using the cumulative average method;
- Comparison of this value toward the basic value average (SN), Lower Warning Limit (LWL), Lower Control Limit (LCL);
- Making conclusion for the student status according the position of the average score.

For of the group performance trend k_{av} is the mean of the score for all students.

As benchmark values in performance trend analysis are used: SN, LWL and LCL. They can have different values but according to the condition: SN>LWL>LCL. LCL is defined as the minimum score for the student to "pass". The analysis if the student score is in the intervals defined by LWL and LCL is essential for diagnostic of "risk" situations, concerning poor performance and possible problems in learning.

Information for the position of k_{av} towards SN, LWL, LCL is the first part of the diagnostic message and it has strong psychological effect for the students, especially for the well motivated.

The difference between the score for the separate tests is a characteristic used for the trend analysis.

$$\Delta_{i,i-1} = k_{i-1} - k_i \qquad (1)$$

where k_i – score for the i^{th} test

k_{i-1} – score for the $(i-1)^{th}$ test.

k_0 is an initial value. It is possible to use the result from the Initial test or SN as a value for k_0. SN is the average score value for the same subject for the three resent academic years.

The suggested approach is a step-by-step analysis, using the score from the current test and previous two tests.

For the purpose of the trend analysis the following divisions are introduced:

Fail – division δ1;

Satisfactory – division δ2;

Good – division δ3;

Very good – division δ4;

Excellent – division δ5.

The trends can have increasing tendency, decreasing tendency, stability around particular evaluation mark or instability.

The suggested approach for the trend analysis is:

If the two differences are positive, $\Delta_{i,i-1}>0$ and $\Delta_{i-1,\,i-2}>0$ – there is increasing tendency;

If the two differences are negative, $\Delta_{i,i-1}<0$ and $\Delta_{i-1,\,i-2}<0$ – there is decreasing tendency;

If the two differences $\Delta_{i-1,\ i-2}$ and $\Delta_{i,i-1}$ have alternative signs, and for the test results is fulfilled:

$$y_{i-2} \in \delta_k$$
$$y_{i-1} \in \delta_l$$
$$y_i \in \delta_m \qquad (2)$$

and

$$|k-l| \geq 2$$

and

$$|l-m| \geq 2 \qquad (3)$$

there is instability;

If the two differences $\Delta_{i-1,\ i-2}$ and $\Delta_{i,i-1}$ have alternative signs, and for the test results is not fulfilled (2) and (3) there is stability around the average value for the least three tests.

All possible combinations for the trend tendencies and the average value are explored at two levels - for individual student and for the student group

Analysis of the time indicators

Using the parameter time it is possible to define some of the main characteristics of the learning process, such as: intensity of learning, uniformity in mastering, delays toward the schedule.

The combination of evaluations of the time for mastering and the level mastering gives a more realistic idea of the current student learning status. These two parameters are often used for evaluation and classification in other areas of life such as sports and competitions.

The planned schedule defines the particular time for i^{-th} test T_i and the planned total time for the course T_g, permissible delay for each test pd, permissible total delay within the course ptd. If the planned time for current test is T_i and the real time for the student is t_i, the difference between two of them Δt_i is used for diagnostic purposes.

$$\Delta t_i = T_i - t_i \quad \text{i=1, 2, …, N} \qquad (4)$$

where N – number of tests

The analysis of deviations from the planned schedule includes:

$\Delta t_i > 0$ - ahead schedule (5)

$\Delta t_i = 0$ - according to schedule. (6)

$\Delta t_i < 0$ - behind schedule (7)

It is considered appropriate to define maximal permitted delay for each test pd (permitted delay) and permissible total delay within the course ptd. The values for pd and ptd are defined for each particular course according to its duration and its specific requirements. The learning parameters uniformity and intensity are connected with the planned schedule and for this reason introducing these permitted delays is a way to provide conditions to fulfill these requirements.

The possible cases that are considered are:

a) $|\Delta t_i| \leq pd$, where pd is the permitted delay. (8)

b) $|\Delta t_i| > pd$, where pd is the permitted delay. (9)

c) Tendency of increasing the delays

$\Delta t_i < 0$ and $|\Delta t_{i-2}| < |\Delta t_{i-1}| < |\Delta t_i|$. (10)

d) The delays ($\Delta t_i < 0$) are above the permitted values and they are lasting (enduring) they lead the student to impossibility to manage within the planned time

$|\Delta t_{i-1}| > pd$ and $|\Delta t_i| > pd$ … ……(11)

d) The sum of the delays ($\Delta t_i < 0$) exceeds the total permissible delay for the course ptd

and $\sum_{i=1}^{nd} |\Delta t_i| \geq ptd$, nd is the number of delays…… ……(12)

It is accepted that the student interrupt his learning if

j-r>MT (13)

where: the number of current test according to the schedule is j; the number of the test that is made by the student is r; MT is number of the missed tests that are permitted.

Analysis of the time indicators gives essential diagnostic information for the progress of learning. It is especially valuable when we combine it with analysis of the other learning characteristics.

Analysis of the test answers

Each test question is characterized with the following characteristics: subject; test number; question number; number of points (difficulty) - 1, 2, 3; how many times is given wrong and/or no answer; type of the question;

weighting coefficient; which subtopic concerns; additional question feature; other features. These characteristics are used in the analysis of the incorrect or unanswered questions.

One of the main characteristic of the test questions for diagnostic purposes is their importance (significance). The importance of these types is given by their weighting coefficients [1].

The analysis of the student answers includes:

- Number (percentage) of the wrong answers for each test.
- Grouping the questions with wrong/no answer by the subtopics they concerns.
- Defining subtopics from the teaching material with more than 50% wrong answers (50% is an exemplary value.
- Showing the type of the questions with wrong answers.
- Percentage of the wrong answers for each type (or information about the total number of the questions from the group and the number of the wrong answers).
- Defining question types with more than 50% wrong answers (dominant types) (50% is an exemplary value).
- Detecting question types with more than 50% wrong answers that repeat for a few tests (repeating dominant types) (50% is an exemplary value).
- Defining questions that are given wrong/no answers 2 and more times (tracing if the "old" questions are answered).
- Defining the mastered material for the whole topic – allows us to make conclusion and change in the input parameter "quantity teaching material" (Fig.1).

The following values are used for this analysis:

QN – number of questions

p_i – theoretical number of points for the i^{th} question

ap_i – answer points for the i^{th} question

WQ_i – weighting coefficient for i^{th} question, according its type

WT – (theoretical) total weight of the test

WA – weight of the student test answers (mastered material)

NM – non mastered material

t1, t2, t3, .. – type of the questions.

$$WT = \sum_{i=1}^{QN} p_i WQ_i \quad (14)$$

$$WA = \sum_{i=1}^{QN} ap_i WQ_i \quad (15)$$

$$NM = WT - WA \quad (16)$$

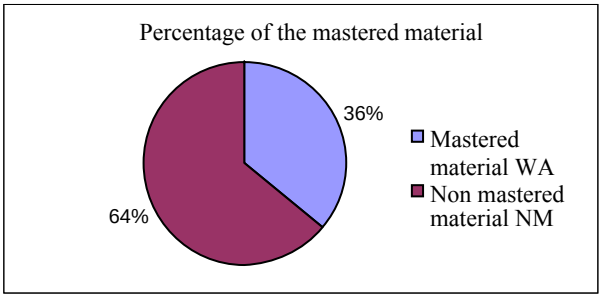


Fig. 1. A theoretical example – percentage of the mastered material

- Defining the mastered material by question types (Fig.2).

The following symbols are used for the calculations:

J – number of question type

p_i – theoretical number of points for the i^{th} question

ap_i – answer points for the i^{th} question

tj , j=1,2,…,11 – question type (t1, …, t11)

QNj– Number of the questions from type tj,

W_j- weighting coefficient for question type tj,

WT_j- theoretical weight for the questions type tj,

WA_j – weight of the student test answers for question type tj, (mastered material for type tj)

WA_j – non mastered material for question type tj

$$WT_j = \sum_{i=1}^{QN_j} p_i W_j \quad , \ j=1,2,\ldots,11 \quad (17)$$

$$WA_j = \sum_{i=1}^{QN_j} ap_i W_j \quad , \ j=1,2,\ldots,11 \quad (18)$$

$$NM_j = WT_j - WA_j \quad (19)$$

The results from analysis in the list positions can be summarized for all conducted tests.

The developed approach for diagnostics is estimated as sufficient minimum for effective FDA and control of the learning process that is universal. The approach is open to add new options, for instance prognosis of the results in

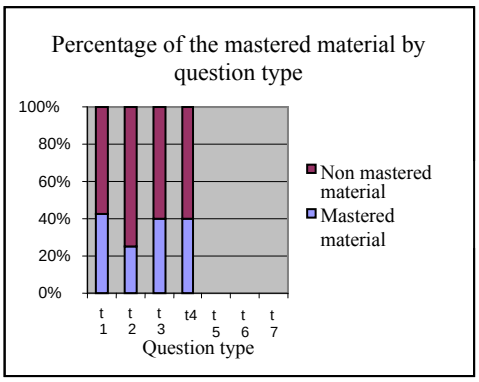


Fig. 2. Percentage of mastered and non mastered material by question type

cases where this feature is seen as useful or to change some of the suggested analysis (calculations) for particular requirements of a teaching course. Introducing specific details will direct the diagnostic to meet particular education needs.

DIAGNOSTIC MESSAGES AND CONTROL DECISIONS

The task of diagnostic assessment is striving to control the learning process with the goal its improvement and optimization (maximal approaching the standard). Control of the learning is to define the approaches and means for influence and their application. In the suggested system the control is performed by generating of diagnostic messages and concomitant control decisions for changing the input parameters of the learning process: quantity teaching material, administrative and psychological impacts.

The study suggests classification of the DM and CD based on the analysed characteristic and in version for the single learner and a student group. The diagnostic message itself already has psychological effect on the student. This effect is not always the desired from the teacher effect. That is why it is necessary to include additionally particular advice, help, praise etc. to direct the effect of the DM into the right (desired) direction and to turn it into a control action (impact). So, in practice, the control action (impact) is combination of diagnostic message for the observed parameters and particular decisions concerning quantity teaching material, administrative and psychological CD. Formulation of the psychological component of the DM can be subject of development for an expert team of psychologists, specialist in pedagogy, specialists in the particular subject area, based on the teaching experience from previous run courses.

Developed cumulative test system for automated formative and diagnostic assessment includes automated generation of DM and CD for most of the cases considered as "common" in learning. For specific situations the teacher has control function. The teacher is a supervisor. He has access to the full information for all students and always can add or change something. The automated analysis is effective assistive tool for the tutor (tutors' team) and reduces to a great extent his pedagogical load in supervising students.

CONCLUSION

Diagnostic assessment has an essential contribution to improve quality of e-learning. There is practical need to incorporate it within university e-learning programs. It is necessary to explore the ways to implement the cumulative test system as part of working e-learning system presenting the tests according to the standards. A research project at European level is a good way to achieve this goal. Our current research is to further develop the suggested version of the test system to achieve adaptable personal e-assessment considering accessibility and usability issues.

REFERENCE

[1] Nacheva-Skopalik L., Defining the weighting coefficients of the different questions types for diagnostic tests, International Scientific Conference "Unitech'03", 20-21.11.03, TU-Gabrovo.

[2] Nacheva-Skopalik L., Skopalik E., Stoyanov S., Statistical approaches for quality of learning control of a student group, International Conference Intelligent Control Systems, Brno University of Technology FEEC, Czech Republic, August 29th –September 11th, 2005.

[3] Nacheva-Skopalik L., An examination of an intelligent cybernetic learning model for formative assessment and diagnostics in open and distance learning, PhD thesis, University of Teesside, UK, 2007.

[4] Nacheva-Skopalik L., Cumulative test system for automated formative and diagnostic assessment, ICL 2010, Hasselt, Belgium, 15-17.09.2010