Set of models for assessing knowledge in distance learning systems

Lidiya Guryanova1[0000-0002-2009-1451], Roman Yatsenko2[0000-0001-7968-6999], Vladyslav Zarzhetskyi3[0000-0003-2500-2414] and Iryna Lytovchenko4
1, 2, 3, 4 Simon Kuznets Kharkiv National University of Economics, Science avenue, 9-a, Kharkiv, 61166, Ukraine
guryanovvalidiya@gmail.com, roman.yatsenko@hneu.net, vladyslav.zarzhetskyi@hneu.net, iryna.lytovchenko@hneu.net

Abstract. The purpose of the study is to build a set of models for assessing knowledge for the purpose of further use in distance learning systems. The existing approaches to the development of models for assessing knowledge are analyzed and a complex of models for assessing knowledge in distance learning systems is built, based on the spaced repetition method and the individual trajectory model. The complex includes blocks of forming a bank of questions, the acquisition and formation of knowledge, assessment of knowledge, improving the quality of a bank of questions. A set of knowledge assessment models based on the chat bot has been implemented to prepare for EIT in mathematics in Telegram. The educational process was tested using a complex of models.

Keywords: model; assessing: knowledge: question: learning.

1 Introduction

Personalization and adaptation of the educational process is an important criterion of the modern educational process [1]. The implementation of these principles is possible with the use of all the modern technical advances in the field of telecommunication technologies and the Internet [2].

The use of interactive technologies in education does not only increase the creative and intellectual potential of students through self-organization, the desire for knowledge, the ability to interact with computer technology and make decisions independently, but also forms a competent specialist with the necessary subject orientation [3, 4]. Adaptive interactive technologies of knowledge acquisition and assessment allow each student to get an objective result in terms of acquired knowledge, to individualize their learning process, to provide self-control [5, 6, 7].

The purpose of education should come from the student, and the content and assessment of results – from the educational system (educational institution, educational materials and teaching staff) [8, 9]. However, the current dynamics of the needs of

Copyright © 2021 for this paper by its authors. Use permitted under Creative Commons License Attribution 4.0 International (CC BY 4.0).
society and the labor market requires the improvement of the educational process using the latest technical and software solutions [10].

The modern education system offers several forms of its acquisition: full-time (day, evening), correspondence training, distance, network, external, family (home), pedagogical patronage, education in the workplace and dual [11]. Unlike classical forms of education, distance learning is most consistent with the current level of development of society – it is carried out using all the latest technical advances in telecommunication technologies and the Internet.

Distance learning means an individualized process of acquiring knowledge, skills, abilities and ways of human cognitive activity, which occurs mainly through the indirect interaction of distant participants in the learning process in a specialized environment that operates on the basis of modern psychological, pedagogical, information and communication technologies [12].

Distance learning system (DLS) is a system (software application) for administration, documentation, tracking, reporting and provision of educational materials with shared access [13].

The advantages of DLS include compatibility, accessibility, reusability, durability, technical capabilities, adaptability, support of different formats of content, relevance, simplicity and adequacy of assessment [14], but there are also disadvantages, such as technical infrastructure requirements, teachers must integrate teaching materials to DLS, the increased workload for teachers is possible [15]. A wide range of functional capabilities of distance learning systems allows to meet the needs of modern educational trends, significantly improve the speed, comfort and quality of learning.

Many works of domestic and foreign scientists are devoted to the problems of introduction of distance learning, for example the works of Y. Babansky [16], V. Bespalko [17], D. Danilov, F. Tovarishchev, A. Nikolaev [18] and V. Ponomarenko, T. Klebanova, R. Yatsenko [19]. Possibilities of using Moodle DLS are analyzed in the works of V. Gavrilenko, V. Popenko, O. Sokulsky, O. Shumeiko [20], V. Sergienko, V. Franchuk, L. Kukhar, O. Halytskiy, P. Mykytenko [21], V. Stepanov, E. Ponomarenko [22], S. Shilo [23]. Approaches to the construction of adaptive knowledge assessment are analyzed in the works of P. Fedoruk [24], L. Zaytseva, N. Prokofiev [25] and K. Navrotska, D. Shtofel, S. Kostyshyn, V. Makogon [26]. However, the analysis of the works of these researchers revealed that the issue of technical development of an adaptive knowledge assessment system in distance learning systems has not been implemented.

2 Set of knowledge assessment models

The study proposes a set of models for assessing knowledge, the general scheme of which is shown in Fig. 1. It has 4 functional blocks. Each unit has its own task in the educational process – the formation of knowledge elements, their acquisition, knowledge assessment and improving the quality of the educational process.
Fig. 1. General scheme of the set of knowledge assessment models.

**Block 1. Question bank formation**

1.1. Methods of knowledge assessment elements formation
1.2. Methods of determination the level of questions difficulty

**Block 2. Knowledge acquisition and formation**

2.1. Model SM-2
2.2. Model SM-4
2.3. Model SM-5

**Block 3. Knowledge assessment**

3.1. Model of the individual path
3.2. Model of determining the assessment results

**Block 4. Improving the question bank quality**

4.1. Methods of updating the questions difficulty level
4.2. Methods of assessment the quality of the question bank elements

---

**Block 1.** The question bank formation contains methods of knowledge assessment elements formation and methods of determination the level of questions difficulty. This stage is characterized by the formation of a general bank of testing questions, which is divided into two samples – for training and for assessment. The formation of knowledge assessment elements is provided in accordance with the objectives of the discipline, but it is worth highlighting the general recommendations:

- the assessment element must be small and simple, it must be aimed at assessing the indivisible unit of knowledge;
- categorization and division by topics allow to improve navigation and combination of assessment goals;
- clear and unambiguous formulation of questions and answers if available;
- use of visualization elements to improve student perception, such as diagrams, tables, specialized tools for drawing formulas, infographics, etc.

Forming separate samples for learning and assessment allows the student to acquire a certain set of knowledge and then assess their abilities using a different set of questions. This approach allows a student to teach and to objectively assess his level of knowledge. The implementation of this approach can be as follows:

- division of the general bank of questions in a certain ratio, for example, 9 to 1 – 90% for training, and 10% for assessment;
• division according to levels of difficulty: easier to learn and more difficult to assess, or the creation of similarly distributed levels of difficulty sets of questions to make the process of assessing knowledge similar to learning.

To determine the levels of the elements difficulty, it is proposed to use the method of log-scaling, which is based on the conversion of the probability of successful completion of the test element in the level of complexity according to the following formula:

\[ d_i = \ln\left(\frac{100 - p_i}{p_i}\right), \]

where \( d_i \) – the level of complexity of the i-th question;
\( p_i \) – the percentage probability of successfully passing the i-th question.

The obtained levels of difficulty will range from -5 (simplest questions) to 5 (most difficult). To make it easier to use the obtained values, you can adjust them to an offset with a value of 5 to get a scale from 0 to 10. This method is designed to fully automate the process of determining the complexity of questions. It allows you to get more accurate levels of difficulty.

The disadvantages of this method include the fact that it does not show itself well in small question banks and is sensitive to lack of statistics. The question bank formation takes place in advance and requires certain statistics regarding the passage of its elements. It is necessary to modify the levels of the question bank elements complexity periodically, getting more and more statistics, for example, after each testing session.

**Block 2.** Knowledge acquisition and formation is realized with the help of models, which are formed by methods of interval repetition of the question bank for training. Spaced repetition techniques allow you to effectively acquire and generate knowledge, maximizing memory stability and minimizing time consumption through optimal repetition intervals.

One of the most important problems of modern learning is the problem of forgetting. It is that soon after learning a new material, we remember only a small part of it. The less likely we are to repeat what we have learned, the faster we will be able to erase new knowledge from our memory. It has long been known that "repetitio est mater studiorum" (Latin: "repetition is the mother of learning"). In other words, the best way to memorize is to repeat what you have learned. However, we may face frustration when we have to repeat old elements while our teachers or supervisors also want us to learn as much new material as possible [27].

Finding time for learning new material and repeating what has been done is a big problem. Usually, we find an intermediate solution. We spend most of the time learning new information, forgetting what we have learned before, and repeating only the material that is necessary for current exams and other situations [28]. The result of this approach to learning is catastrophic. The result of this approach to learning is catastrophic. Most of the time is wasted, as much of what we learn we forget. Of course, we improve the overall understanding of the material studied, but the understanding is also based on memories and is equally unstable. And it is only a matter of time before we lose most of our investment in education.
The current load does not allow us to fully repeat what we have learned before. Educational systems around the world penalize those who do not learn new material. We get into a ridiculous situation: we quickly study new material, pass certification (test, exam), and then begin to study the next, forgetting the previous one [29]. The solution to this problem can be spaced repetition, which significantly reduces the time required to repeat the studied material. This should solve a significant number of learning problems.

Spaced repetition is a learning method based on the calculation of optimal intervals, which should separate the learning of each individual element of knowledge to ensure a high level of memory retention [30].

The optimal interval is the ideal period of time to separate the acquisition of knowledge. It is used to maximize the memory effect. In practical use, knowledge retention should be the main criterion for optimization. In most cases, the term “optimal interval” means the intervals as a result of which the forgetting index reaches 10% [31].

Optimal intervals are calculated on the base of two conflicting criteria:

- the intervals should be as long as possible to obtain the minimum repetition rate and the best way to use the so-called interval effect, which means that the intervals between repetitions should reach a certain limit to achieve the strongest memorization;
- intervals should be short to ensure that knowledge is still remembered.

In practice, these two criteria become the following: during the study, the intervals should be the amount of time that is necessary for the selected small amount of knowledge to be forgotten. This proportion, called the forgetting index, can range from 3% (for slower and more careful memorization) to 20% (for faster learning, which is characterized by a lower level of knowledge retention).

The forgetting index – is the percentage of knowledge that will not be projected to the memory in the repetition cycle. That is, if the forgetting index is 10%, it can be assumed that 90% of the materials that have been passed during the learning cycle will be kept in memory. This index is used to form the daily cycle of assimilation the knowledge elements, what means that just such part of knowledge may not be remembered during the cycle of assimilation [32].

This approach can help to improve the learning process significantly by reducing time and choosing the optimal intervals for returning to the learned material, in order to minimize the consequences of the problem of forgetting.

Acquisition and formation of knowledge in Block 2 of the set of models is realized with the help of models based on SM-algorithms of spaced repetition [33]. One of the SM-2, SM-4 and SM-5 models is assigned to the student for the entire period of studying. The selection of model is random. This approach was chosen to test each model in practice to identify the best one. A general scheme of training using the method of spaced repetition, which is the basis of all SM models is shown in Fig. 2.
The stage of formation the question set for the educational session, which is the same for all SM models is shown in Fig. 3. It aims at filling the set of questions of a certain length with the elements to repeat or new to study. The size of the sequence is determined by the teacher; it is desirable to set this parameter in the amount of 15-20 elements so that students can effectively conduct an educational session. The set of questions is first filled with elements for repetition, and if there are none, with the new ones. The search for a repeating question is performed according to the value of the repeating interval, which was determined after processing a new question. The number of formation and processing of educational sessions is unlimited.
Questions are issued and passed in any convenient way for the teacher and the student. However, it should be noted that the automation of the entire educational process allows not only not to worry about calculating the optimal intervals of repetition and storage of supporting information for learning, but also to spend more time to improve the quality of content – a question bank.

**Block 3.** Assessment of knowledge is built using the model of individual path and the model of determining the assessment results. An important stage of learning is the assessment of acquired knowledge, skills and abilities of the learner. There are two approaches to this – the traditional, historically developed, which is characterized by a high probability of the presence of a subjective point of view of the teacher and significant resource intensity, and modern, which replaced the first with the development of information technology.
The modern approach uses methods of knowledge assessment that increase the objectivity of testing and assessment of studying results. In order to minimize material and time resources in the educational process adaptive method of knowledge assessment is actively used. This method is characterized by gradual adaptation to the level of the student, which allows to assess his progress adequately and eliminate psychological barriers and problems that arise during learning.

The structure of adaptive knowledge assessment, built on the model of adaptive control of knowledge is shown in Fig.4 [34]. It describes the cyclical process of assessing a student’s knowledge by giving him a series of tasks. The answer to the previous task affects the next – thus there is a gradual adaptation to the level of knowledge of the student.

![Fig. 4. Structure of adaptive knowledge assessment.](image)

The mechanism of adaptive knowledge assessment is aimed at optimal use of student resources and assessment system. This approach is characterized by significant preparation to the testing process by the teacher (preparation of the question bank, determining the quality of elements, the formation of evaluation criteria, etc.), but determining the level of student’s knowledge is more effective. It should be noted that the adequacy of the obtained results significantly depends on the quality of preparation to the test.

The selection of the next question for this method of assessment is based on the levels of the elements of the question bank complexity and the student's abilities. The level of difficulty of the question can be interpreted as follows: example, we have an element with a level of difficulty of 3, each student who has a level of ability 3 has 50% to work successfully, and the higher the level of abilities he has, more easily he can deal with it and vice versa.
The trajectory of testing according to this principle is shown in Fig. 5, where the test questions are on the abscissa axis and the level of difficulty is on the ordinate axis. As you can see, Fig. 5 also shows the level of the student's abilities, reaching which more and more rapidly he begins to oscillate around. That is, the student reaches the level of complexity of the questions with which he has 50% success. This bar characterizes the level of student’s abilities.

Fig. 5. The trajectory of testing using levels of questions’ difficulty.

This approach allows the student to reach the questions of his level quickly and deal with them comfortably, without getting questions which are too easy or too difficult for him. Basing on such a mechanism, it is possible to implement adaptive models of knowledge assessment.

Models of individual trajectory and determination of assessment results a modified computer testing algorithm which is based on the Simpler CAT Algorithm, authored by Wright, B.D. and consists of several stages [35].

Stage 1. Initialization of the testing process. Setting variables in relation to passing the test according to the formula (1) and the following auxiliary variables at the discretion of the teacher:

- sufficient level of difficulty to pass the test (T);
- minimum level of estimation error (S_{min});
- minimum (L_{min}) and maximum (L_{max}) number of test questions.
Stage 2. Preparation and issuance of questions to the student. Search for a question close to the current difficulty, taking into account the direction in accordance with the correct answer to the previous question by formulas (2-3), updating the current difficulty of testing by formula (4) and issuing a question to the student.

\[
\begin{align*}
\text{correct answer: } D' & \geq D, \\
\text{incorrect answer: } D' & \leq D,
\end{align*}
\]

where \( D' \) – the difficulty of the next question from the question bank;
\( D \) – the current difficulty of the question.

\[ D = D', \]

where \( D \) – the current difficulty of the question;
\( D' \) – the difficulty of the question received from the question bank.

Stage 3. Elaboration of the student's answer to the test question. Modification of the main variables for testing according to formulas (5-6), as well as if the student's answer is correct, then according to formulas (7-8), otherwise (9-10).

\[
\begin{align*}
H &= H + D, \\
L &= L + 1,
\end{align*}
\]

where \( H \) – total difficulty of processed questions;
\( D \) – the difficulty level of the processed question;
\( L \) – the number of processed questions.

\[
\begin{align*}
R &= R + 1, \\
D &= D + \frac{2}{L'}
\end{align*}
\]

where \( R \) – the number of correct answers;
\( D \) – current level of difficulty;
\( L' \) – the number of processed questions.

\[
\begin{align*}
W &= W + 1, \\
D &= D - \frac{2}{L'}
\end{align*}
\]

where \( W \) – the number of incorrect answers;
\( D \) – current level of difficulty;
\( L \) – the number of processed questions.
Step 4. Determining the assessment of the achieved difficulty level of testing questions and assessment errors. If all the answers are correct, then formulas (11-12) are used, if all the answers are incorrect, then use formulas (13-14), otherwise use formulas (15-16).

\[ B = \frac{H}{L} + \ln\left(\frac{R}{0.5} - 1\right), \quad (11) \]
\[ S = \sqrt{\frac{L}{0.5(1 - 0.5)}} \quad (12) \]

where \( B \) – assessment of the achieved difficulty level of testing questions;
\( H \) – the total difficulty of the processed questions;
\( L \) – the number of processed questions;
\( R \) – the number of correct answers;
\( S \) – assessment error.

\[ B = \frac{H}{L} + \ln\left(\frac{0.5}{W - 0.5}\right), \quad (13) \]
\[ S = \sqrt{\frac{L}{0.5(W - 0.5)}} \quad (14) \]

where \( B \) – assessment of the achieved difficulty level of testing questions;
\( H \) – the total difficulty of the processed questions;
\( L \) – the number of processed questions;
\( W \) – the number of incorrect answers;
\( S \) – assessment error.

\[ B = \frac{H}{L} + \ln\left(\frac{R}{W}\right), \quad (15) \]
\[ S = \sqrt{\frac{L}{R \times W}} \quad (16) \]

where \( B \) – assessment of the achieved difficulty level of testing questions;
\( H \) – the total difficulty of the processed questions;
\( L \) – the number of processed questions;
\( R \) – the number of correct answers;
\( W \) – the number of incorrect answers;
\( S \) – assessment error.

Stage 5. Making the decision on finishing the test. This stage occurs according to one of the following criteria:

- if the question bank is exhausted, go to step 6;
- if the maximum number of questions is completed, go to step 6;
- if the minimum estimation error is satisfied, go to step 6;
- if, after passing the minimum number of questions, all the answers are correct or incorrect, go to step 6;
- if the minimum number of questions is passed, the transition to stage 6 can be made (the decision is made by the student);
if the student demonstrates non-standard behavior of passing the test, the transition to stage 6 can be made (the decision is made by the teacher);
• otherwise go to stage 2.

Stage 6. Formation and issuance of the test results. Determining the rate of correct answers according to formula (17), as well as testing scores - according to formula (18). The verdict is also formed according to formulas (19-20): if inequality (19) is satisfied, then testing is passed, if (20) is not passed, otherwise we have a zone of uncertainty.

\[ C = \frac{R}{L}, \]  

(17)

where

- \( C \) – the rate of correct answers;
- \( R \) – the number of correct answers;
- \( L \) – the number of processed questions.

\[ A = \frac{\sum C_i D_i}{\sum D_i}, \]  

(18)

where

- \( D_i \) – the difficulty of the i-th passed question;
- \( C_i \) – correctness of the answer to the i-th passed question, at the correct answer acquires value 1, at incorrect – 0.

\[ B - S > T, \]  

(19)

\[ B + S < T, \]  

(20)

where

- \( B \) – assessment of the achieved difficulty level of testing questions;
- \( S \) – assessment error;
- \( T \) – a sufficient level of difficulty to pass the test.

The process of assessing knowledge using the model of individual path is aimed at determining the optimal value of the level of assimilation of material passed by the student. Optimality is the rational use of student resources and assessment system. As a result, we have the level of knowledge of the student in terms of the achieved level of complexity, assessment error, which indicates the quality of testing, and statistics of student learning activity, which can be used to further improve both the question bank and knowledge assessment system.

The last block of the set of knowledge assessment models (Block 4) deals with the improvement of the quality of the question bank. The components of this block include methods for updating levels of difficulty and methods for assessing the quality of the elements of the question bank. The first includes methods for determining the levels of difficulty of the test bank elements, but their use has been modified to adjust the levels of difficulty after a sufficient number of knowledge assessment cycles. It should be noted that these methods can be both auxiliary for the teacher for analytical analysis, and independent, provided the automation of determining the level of questions difficulty.
Assessing the quality of the question bank elements is to determine the correctness of the composition of the elements and identify anomalies. The formed element of testing can be considered correct for its correct and unambiguous interpretation by students. Anomalies can include various errors in the assembly of the test item, which do not allow to answer it correctly. The in-time detection of problems of the question bank allows to maintain a sufficient level of adequacy of the knowledge assessment process and to guarantee the successful acquisition of knowledge by students.

Thus, a set of models for assessing knowledge is aimed at full coverage of the process of acquisition and assessment during training: the formation of knowledge elements, their acquisition, assessing the level of material assimilation and improving the learning process, using feedback. All this allows to improve the quality of the educational process.

3 Approbation in Telegram messenger

The set of knowledge assessment models was implemented as a chatbot in the Telegram messenger @HNEU_ZNO_math_bot, using the Python programming language and the SQLite database. The chatbot has two modes of operation: learning (default) and knowledge assessment.

The question bank was formed on the basis of tasks for preparation for ZNO in mathematics. Statistics on testing for this bank of questions were taken from the competition “ZNO Mathematics: BOT Challenge” [36]. (about 670 thousand passes from almost 7 thousand users, among whom the vast majority are students of graduating classes).

Determination of difficulty levels was carried out by the share of correct answers of the first attempt to process the question, it means that only the first answers of each user were taken into account. Thus, their levels of complexity were formed using the method of log-scaling with offset. The question bank has 500 elements and difficulty levels are distributed from 2.65 to 7.71 on a scale from 0 to 10. Most levels of difficulty range from 3.11 to 5.87, so we can assume that the question bank is filled with elements of medium difficulty. A histogram of the distribution of difficulty levels is shown in Fig.6.
Fig. 6. The histogram of the difficulty levels distribution.

The approbation started on November 28, 2019 at 18:43 and ended on December 3, 2019 at 22:39. A total number of 318 users took part in the bot and passed 3210 tasks. 179 users took part in the training process and passed 1463 tasks, 1149 of which were successfully mastered at the end of the approbation. The average number of completed tasks by the user is 8.17, and successfully mastered – 6.42. Mostly from 0 to 14 tasks were mastered. Several users have indicators of more than 50 learning elements.

The distribution of the average number of students passing questions before mastering it at the current time is shown in Fig. 7. The majority of students worked on a question with an interval of 1.19 to 1.8 times on average. For example, the average number of passes for one of the active users was 1.48. The lower this indicator is, the less time-consuming is the learning process, but it should be noted that it tends to increase until the user has mastered the element at a sufficient level.
The knowledge assessment mode was used by 35 users, completing 46 assessment sessions. The following statistics were calculated for completed assessments lasting from 2 minutes to an hour to weed out inadequate testing. The number of passed tests is 23, the average length of testing is 12.65 questions, and the duration is 20 minutes 47 seconds. The average proportion of correct answers is 0.54, the average level of achieved difficulty is 4.71, and the assessment error is 0.59. The average score for testing is 0.5 out of 1.

4 Conclusions

Thus, the proposed set of models for assessing knowledge provides a full-fledged distance learning environment where students can both learn and assess their level of knowledge. Approbation confirmed the effectiveness of the learning process, showing a small share of user costs for the repetition of elements, what means that more effort was given to the acquisition of new knowledge.

The average level of knowledge assessment results is related to the lack of motivation of users and lack of time for a sufficient level of mastery of the question bank. Further training by the method of spaced repetition will improve user’s the level of knowledge, and the motivation for this process may be future competitions based on a chatbot to prepare for ZNO in mathematics and other disciplines.

References

21. Sergienko V.P., Franchuk V.M., Kuhar L.O., Galitskiy O.V., Mikitenko P.V.: Methodical recommendations for the establishment of tests for the control system of educational materials MOODLE 2.5.x [in Ukrainian], NPDU, 100 p. (2014).


