Extension of Electronic Testing Systems on the Example of Testing SQL-Queries

Tatyana S. Karpova
Department of Mathematics and modeling, Emperor Alexander I St. Petersburg State Transport University
St. Petersburg, Russia
t.s.karpova@gmail.com

Svetlana Yu. Malysheva
Department Mathematics and modeling, Emperor Alexander I St. Petersburg State Transport University
St. Petersburg, Russia
SvetlanaMalisheva315@yandex.ru

Nikolay N. Teslya
Laboratory of computer aided integrated systems, St.Petersburg Institute for Informatics and Automation of the RAS
Saint Petersburg, Russia
teslya@iias.spb.su

Abstract

The presence of a system of distance learning (education), abbreviated as SDO, is currently a mandatory component of the educational environment of any University of the Russian Federation, accredited for the right to conduct educational activities in accordance with State standards.

1 Introduction

On the recommendation of UNESCO in most publications remains the international name of such information environments, distance learning. without translation [Ser12]. Today's generation lives next door in a computer environment almost around the clock, so the use of information and communication technologies for learning is natural and necessary for him. Russian Universities have to urgently structure, move away from outdated technologies used in the educational process and do it as quickly as possible. It is no secret that most modern students prefer not to attend lectures, because their material can be easily read from a smartphone or laptop in a much more comfortable environment than a lecture hall of the University. Our state adequately responds to the changed conditions and in the new educational standards everywhere demand to reduce the volume of lecture material to a minimum, leaving the discipline of practical training and laboratory work.

Most foreign researchers perceive e-learning as an educational paradigm. They define e-learning as "an innovative learning approach applied to provide a well-designed interactive learning environment to any learner, anywhere and at any time, using the resources of various digital technologies along with other forms of learning materials suitable for an open learning environment. E-learning is making the transition from a data management system to a knowledge management system" [Ser12].

The transition to a new stage of economic development – to the digital economy is also consonant with the active use of e-learning systems in the educational process. However, a simple transfer of printed materials, textbooks and teaching AIDS in the electronic educational environment will not give a clear effect of the assimilation of knowledge, which are set out in this electronic content. Yes information is provided, the knowledge stored in textbooks is publicly available, but there is no guarantee that this knowledge, expressed in electronic form, will be perceived and assimilated by students better than in boring lectures. But the time for presentation of the material is limited by modern standards, the volume of classroom work of teachers at the same time only increased and the load on the teaching staff increased. Feedback components are used to assess the quality of the acquired knowledge in modern electronic learning environments. In most environments, these components include the ability to provide any reports, abstracts, solved specific tasks as well for verification in electronic form. However, such a decision only aggregates the situation, the amount of materials that require individual verification by the teacher increases and this puts the teacher before a choice: either to reduce the number of practical tasks, or to abandon their thorough verification. Both options reduce the quality of learning-without constant, regular and very careful feedback there can be no guarantee of

In: B. V. Sokolov, A. D. Khomonenko, A. A. Bliudov (eds.): Selected Papers of the Workshop Computer Science and Engineering in the framework of the 5 th International Scientific-Methodical Conference "Problems of Mathematical and Natural-Scientific Training in Engineering Education", St.-Petersburg, Russia, 8–9 November, 2018, published at http://ceur-ws.org

mastering the theoretical material that is presented in the content.

The only way out in this case is to build an adaptive testing system that will automatically check everything that is possible and free the teacher from the routine work of checking the initial basic concepts of the taght field of knowledge. A set of essential question Bank and the preparation of these questions themselves also requires a significant amount of time and effort from teachers, but there is a real prospect: once typed test questions can be used repeatedly. In addition, the completion and expansion of the Bank of questions and competitive construction of test models will eliminate empty guessing and provide an effective process of self-education for students, which we strike for.

However, with all the variety of types of test questions in most areas of knowledge, there are professional problems or tasks that can not be implemented available tests. These tasks can have many possible solutions and it is often impossible to foresee and list them all. In this case, a different concept of feedback is required. The authors propose to use the concept, in which there are no attempts to simulate the process of solution, and the student is given the opportunity to solve the problem-the problem in any way and check the result. For this purpose, it is assumed that the e-learning system has a mechanism for the implementation of the solution proposed by the student on the given source data and the mechanism for the implementation of the solution proposed by the teacher on the same data, and the evaluation of the solution is carried out by comparing the results. If the results are the same, you can consider the solution correct, otherwise – no.

Of course, for different classes of problems in different disciples should be supported by different "performers" developed algorithms for solving problems.

To implement this task, the authors choose the task of getting the skills of writing semantically correct SQL queries. From many years of teaching experience, the following result was experimentally obtained: if a student on 10 different databases makes no more than 3 errors when executing 100 SQL queries, then it can be guaranteed that he will make no more than 3% of incorrect queries on any other databases. The authors do not take into account the syntax – it is checked by all translators and mastered quickly enough. Here, the checks are performed exactly within the meaning of the SQL query.

The teacher spends from 5 to 10 minutes to check one SQL query of average complexity. And if, as defined above, for a steady skill of correct writing on SQL-queries is required to check at least 100 on SQL-queries of each student, the elementary calculation shows that the teacher requires a group of students of 25 people almost 300 full hours of study time, which, of course, almost impossible. In modern information systems in the analysis of information skills correct formation of arbitrary SQL-queries are required constantly. Therefore, almost all interviews are checked for the presence of this skill.

2 The Skills

Figure 1 shows the order of formation of skills in any problem domain.

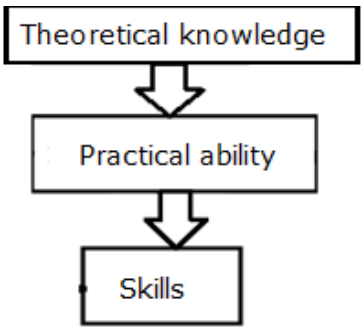

Figure 1: Hierarchy of professional competences

Knowledge is formed in lectures and reading textbooks, practical skills-in the process of practical and laboratory work, and skills are formed with experience, ie, with multiple assignments.

In order for students to be able to apply the theoretical knowledge and practical skills acquired during their studies at the University, they must move into skills. The problem-solving skills will make students competitive.

The authors set a goal-to develop a system of automatic check for SQL-queries. In this case, students can independently develop the skills of correct formation and SQL-queries and confirm it during the control testing. However, to do this, the system must be able to replenish the test databases, fill them with specially verified data, allowing in any situation to identify a semantic error.

3 The Analysis Of Systems Performing The Verification By SQL Queries

In the public domain, there are ready-made solutions that allow you to check the syntax written in the SQL-query [Jew11]. Also, the check can be performed by comparing with the text of the correct SQL query stored in the database (DB), without its execution. This approach has the disadvantage of not taking into account the sequence of data from multiple tables.

From the point of view of the system analyst, they are not suitable for the task, as they are aimed solely at checking the syntax compiled by the SQL query, which is necessary, but does not contribute

to training. Therefore, it was decided to develop its own testing methodology.

4 Description Of The Proposed Mechanism

The complexity of automatic verification of the student's response and the correct SQL-query teacher is the existence of different ways of writing the correct SQL-query to the given task, that is, the translation of the semantic question in the standard SQL-query is not formalized and is creative. For example, a semantically correct SQL query can have multiple ways of writing because of the different order of column or table names used in the query.

During the analysis of possible methods of semantic analysis of SQL-queries, the procedure of comparing the results of the correct SQL-query prepared by the teacher and the results of the SQL-query written by the student was proposed [Kar18].

Figure 2 illustrates the process of comparing the results of a teacher's and a student's SQL queries. The architecture of checking the results of SQL-query allows you to check different versions of its writing.

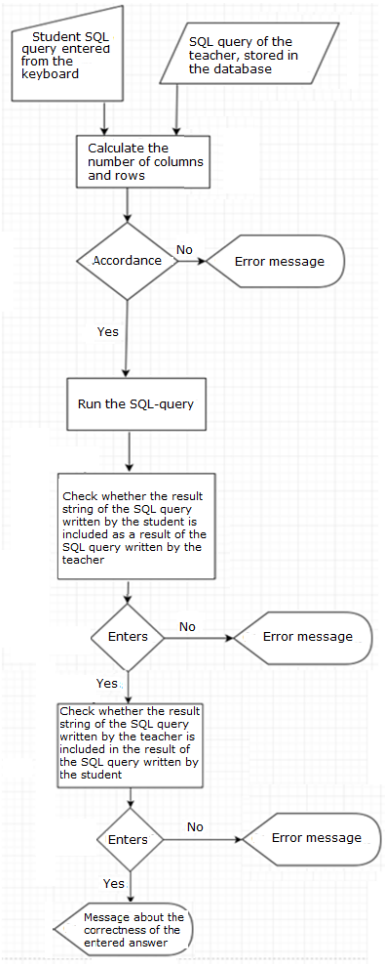

Figure 2: Algorithm of check of conformity of test results

The development of a methodology for comparing and checking the answers of the student and the teacher included the implementation of some operations of relative algebra in a limited version of the language with SQL and the formation of the algorithm shown in the form of a flowchart in figure 3.

When building a system of automatic testing of query checking in SQL, the authors faced the problem of the absence in the free version of the database management system (DBMS) MySQL the ability to execute a query directly corresponding to the difference in relations. This limitation applies to the considered DBMS, which is the basic in the system of support of the educational process e-leaning Moodle. Moodle [Moo19] – open-source system implemented in PHP, it was decided to develop a mechanism for reflecting the difference operation of relative algebra into the difference operation over associative arrays of PHP language to implementation this functionality [Kar18].

During the request processing SQL queries on the DBMS server MySQL using PHP language, you are working with the library of functions that allow you to perform queries of the SQL language in database. The result of the query is an associative array that actually models a two – dimensional table where column area attributes and its rows are the values of the result tuples.

However, among the common operations on associative arrays there is no subtraction operation for multi-dimensional arrays. This operation is available only for one-dimensional linear array.

The General mechanism for checking the correctness of SQL queries prepared by students is shown in figure 3.

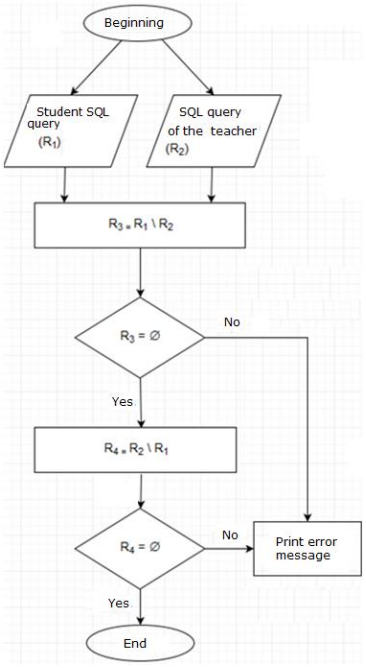

Figure 3: The developed algorithm for comparing the difference with the implementation of the

relative algebra operation, used in the training system

As can be seen from figure 3, this algorithm uses the relative algebra operation – the difference of relations.

The difference between the relations P1 and P2 is the set, which includes a set of tables belonging to R1 and not belonging to R2, and:

$$R_3 = R_1 \setminus R_2 = \{r \mid r \in R_1 \wedge r \notin R_2\}. \quad (1)$$

To implement the algorithm in PHP, you first need to check whether the schemes are equivalent and only then proceed to the execution of the difference operation. This is initially done by comparing the ranks of relationship names and attributes. In this case, the algorithm for performing the operation of the relative algebra of difference is as follows, as shown in figure 4.

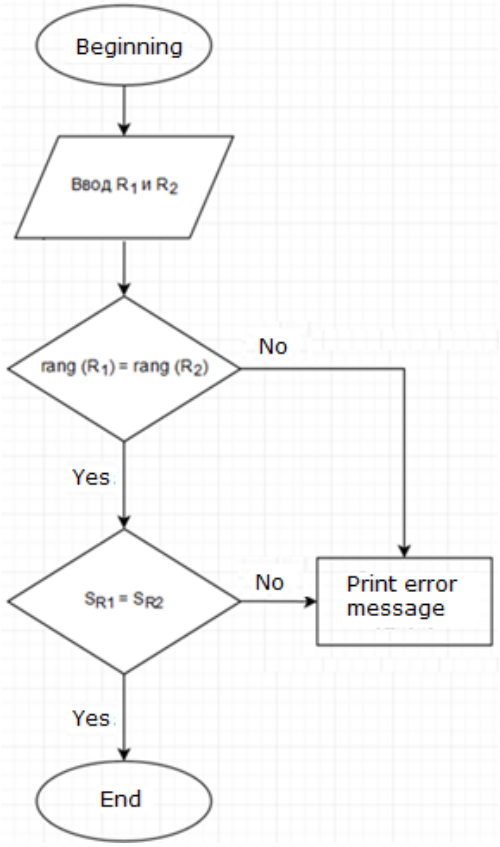

Figure 4: Checking the equivalence of the compared operation schemes

5 Design And Creation Database

To implement the above algorithm in the process, it was decided to create a metadata database at the first step, then create a set of training databases [Nev06] (figure 5).

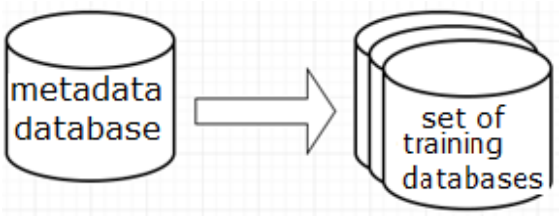

Figure 5: Types of database

Based on the description of the subject area, in the process of communication with experts in this field, the need for the following entities was formulated:

- db,
- tables,
- stolb,
- zapros.

An important stage of the project was the design of the database using CASE (Computer Aided Software Engineering)-system. Modern CASE-systems — means of development of not only software systems, but also organizational and management. The purpose CASE-средств – to separate processes from the design of programming processes.

To ensure the independence of the design of the logical structure of the database was used modern free-distributable CASE-система OpenSystemArchitekt. The choice of this product is justified by its correctness in the formation of the classical infological model EntityRelationShip – ER, shown in figure 6, with the possibility of obtaining a database generation script in accordance with the standard SQL92.

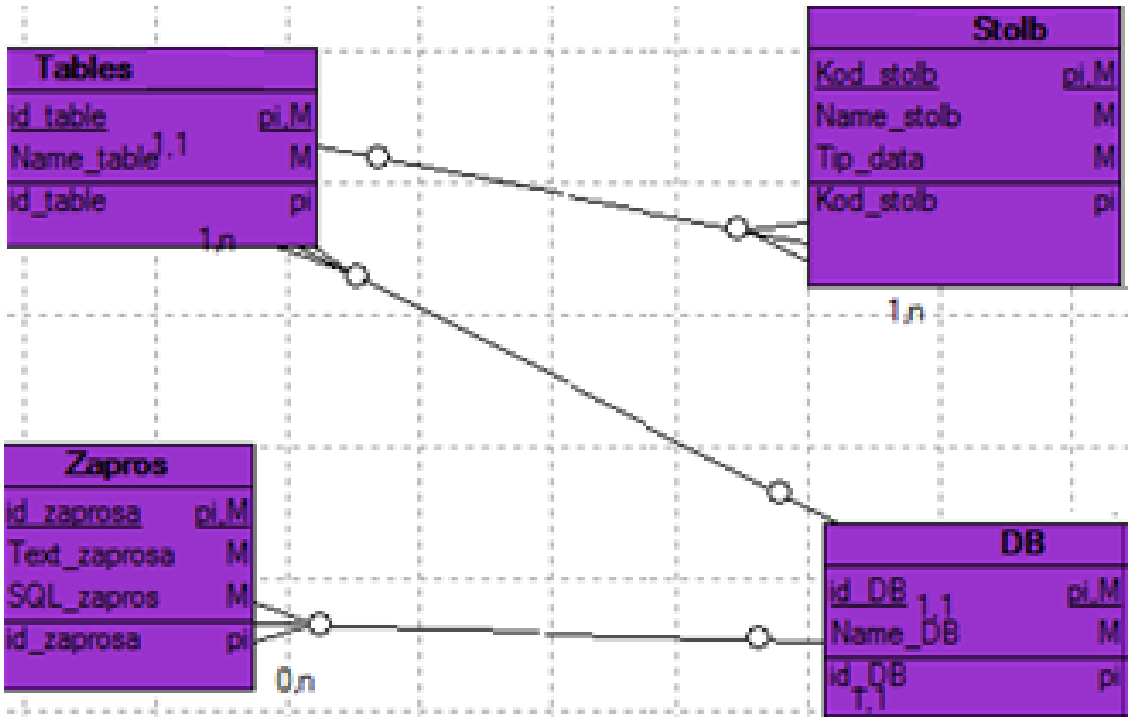

Figure 6:ER model database

Then, based on the logical model, a physical model was created, as shown in figure 7.

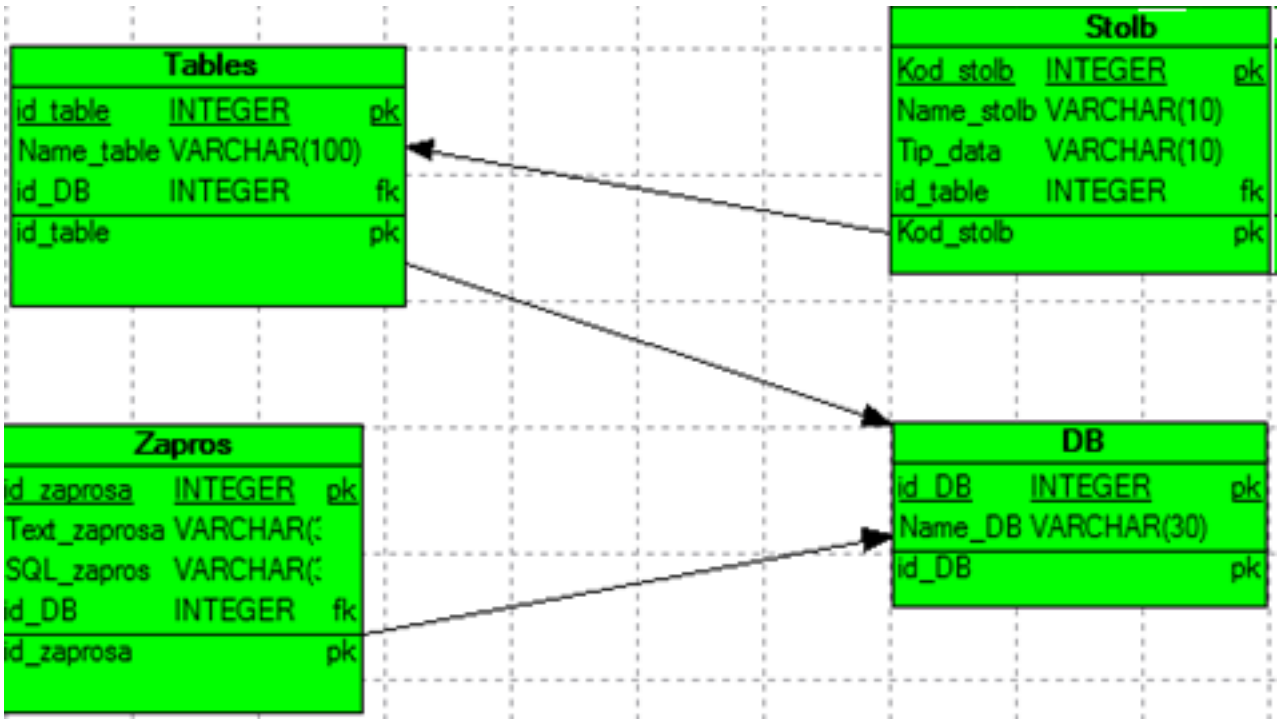

Figure 7: Physical database model

Then the script of database generation was received, imported to Denwer-local server with PHP, Apache and MySQL support, allowing to work with the site code without connecting to the Internet.

In the table "db" the data is stored on the created training databases:

– serial number assigned to automatically,
– the name of the database,

– brief description of the subject area of the database.

Table «tables», the structure of which is shown in figure 9, contains data on the list of all tables included in all training databases, namely:

– serial number assigned to automatically,
– the name of the table,
– database number to which the table belongs,
– number of columns in the table,
– name of the table in English.

In table «stolb» data is stored on the columns of all database tables, indicating the data type:

– serial number assigned to automatically,
– column name,
– data type,
– the table number refers to the column

Table «zapros» stores information about the SQL query environment:

– serial number assigned to automatically,
– query text,
– the correct answer in the form on a SQL query,
– database number to which the request is written.

To create a training database "Products" developed at the initial stage of the database tables, which includes the following tables:

- db,
- tables,
- stolb,
- zapros,

new will be added :

- tovar,
- zakaz,
- klient,
- postav.

6 Implementation Of The Methodology, Functional Development

The implemented technique assumes storage in the database of the list:

• training database,
• table of the training database with the description,
• fields with data type and meaning,

that was implemented and described above. When implementing the algorithm illustrated in figure 4 in PHP, it is sufficient to use the mysql_num_fields () function to compare ranks, which returns the number of elements in the associative array of the result, and to perform the schema comparison operation, you need to perform the function of comparing keys of associative arrays using the array_diff_key () function [Kar18]. The use of functions is shown in the program code in figure 13.

zapros.php

```php
<?php
include('common.php');
include('func.php');
function zapros($stud,$prep) {
$result_1 = mysql_query($stud);
$c_1=mysql_num_fields($result_1);
$n_1 = mysql_num_rows($result_1);
$result_2 = mysql_query($prep);
$c_2=mysql_num_fields($result_2);
$n_2 = mysql_num_rows($result_2);
IF (($c_1<>$c_2) or ($n_1<>$n_2)) return 0;
$r=array_diff_key($result_1, $result_2);
If (srav($result_1,$result_2,$n_1)==1) return
else return 0;
}
?>
```

Figure 13: Source code comparison of the ranks of names of relations and attributes

If, when checking the algorithm shown in figure 4, the scheme of operations of the equivalents to be compared, it is necessary to proceed to the subtraction operation, the implementation of which is shown in figure 3.

In this case, in the absence of the possibility to carry out the subtraction operation directly on the multidimensional associative arrays, a special function srav (a1, a2), was developed, presented in figure 14.

php

```php
<?php
function srav($stud,$prep,$n_1)
{
    $n=0; $i=0;
// $stud - result of the request of the student
// $prep - the result of the request of the teacher
While ($row_2=mysql_fetch_row($prep))
{
$a[$i]=$row_2;$i++;
}
While ($row 1=mysql fetch row($stud))
{
for ($j=0;$j<$i;$j++)
{   $r=array_diff($row_1,$a[$j]);
    if (count($r)==0)$n++;
}
}
if ($n_1==$n) return 1;
else return 0;
}
?>
```

Figure 14: Code to perform the subtraction operation

7 Development And Writing Interfaces

For the formation of training databases, for further testing of students, it was necessary to develop an interface of the teacher, with which he will fill the database metadata and training databases for the execution of SQL-queries.

Using the PHP programming language and the HTML hypertext markup language, the interface was developed for the teacher.

8 Further Development

In the development of the issue raised, the authors plan to integrate the mechanism in the open-source wrapper for the class systems e-leaning Moodle. According to the statistics collected at the summit of the developers of e-courses, we can say that Moodle is the most used in the educational environment.

References

[Ser12] A. G. Sergeev. Introduction to e-learning: monograph of Sergeyev, I. E. Zhigalov, V. V. Balandina; Vladimir. state University named after Alexander Grigorievich and Nikolai Grigorievich Stoletovs. - Vladimir: publishing house of VlSU, 2012. - 182 p. ISBN 978-5-9984-0268-5.

[Jew**11]** B. Nevarez. Inside the SQL Server Query Optimizer, 2011. — 265 p.

[Kar18] T. S. Karpov, S. Yu. Malysheva. System approach to the development of training system for obtaining skills // VI scientific - practical conference with international participation "science of present and future" for students, postgraduates and young scientists. Proceedings of the conference. SPb.: Publishing house Etu "LETI", 2018. 144 - 146 p.

[Moo19] Moodle Docs. [Electronic resource]. – URL: https://docs.moodle.org/36/en/Main_page . – (Date of appeal: 17.02.2019).

[Kar18] T. S. Karpov, S. Yu. Malysheva. Implementation of some relational algebra operations in a limited version of SQL // Problems of mathematical and natural science training in engineering education. Sat. proceedings of the 5th international scientific-methodical conference. 8 - November 9, 2018, St. Petersburg/ ed. – St. Petersburg: PGUPS, 2018. P. 141– 148.

[Nev06] T. Jewett Database Design With UML and SQL. –Department of Computer Engineering and Computer Science California State University, Long Beach, 2006.