PROCEEDINGS

of the Union of Scientists - Ruse

Book 5

Mathematics, Informatics and Physics

Volume 7, 2010

RUSE

ISSN 1311-9184

The Ruse Branch of the Union of Scientists in Bulgaria was founded in 1956. Its first Chairman was Prof. Stoyan Petrov. He was followed by Prof. Trifon Georgiev, Prof. Kolyo Vasilev, Prof. Georgi Popov, Prof. Mityo Kanev, Assoc. Prof. Boris Borisov, Prof. Emil Marinov. The individual members number nearly 300 recognized scientists from Ruse, organized in 13 scientific sections. There are several collective members too – organizations and companies from Ruse, known for their success in the field of science and higher education, or their applied research activities. The activities of the Union of Scientists – Ruse are numerous: scientific, educational and other humanitarian events directly related to hot issues in the development of Ruse region, including its infrastructure, environment, history and future development; commitment to the development of the scientific organizations in Ruse, the professional development and growth of the scientists and the protection of their individual rights.

The Union of Scientists – Ruse (US – Ruse) organizes publishing of scientific and popular informative literature, and since 1998 – the "Proceedings of the Union of Scientists- Ruse".

SERIES 5

"MATHEMATICS, INFORMATICS AND PHYSICS"

VOLUME 7

CONTENTS

CONVENTIONAL KNOWLEDGE TESTING IN COMPARISON WITH INTELLIGENT TEST IN ALGORITHM AREA – AN EXPERIMENTAL STUDY

Galina Atanasova

Angel Kanchev University of Ruse

Abstract: *The basic idea of this paper is to present the result of an experimental study, which exams an intelligent test and a conventional multiple choice test in the algorithm learning area. It is described the main hypothesis proof that the intelligent test covers more didactic goals, range more fully the taught material and leave the learners' better possibilities to show acquired skills in algorithms. The paper focuses on the following basic points: goal, object and methodology of the study, experimental data, test quality evaluation, correlation analysis, and both test comparison.*

Keywords: *computer science, algorithms, task-oriented environments, intelligent testing, experimental study.*

INTRODUCTION

Algorithms are abstract procedural knowledge units, modeling the solution of a given class of problems in different domains such as mathematics, informatics, economics, management, and so on [4]. The deep algorithm knowledge is an important precondition for successful teaching and learning in computer sciences area. There are different teaching approaches such as conventional, part or full computer aided. Many experiment results prove the advantages of the algorithm animation systems (AAS). The combination of algorithm animation with tutor instructions demonstrates most significant benefits [2].

The developers have shifted to build on-line AAS with the appearance of INTERNET technologies. The main reasons are their advantages of platform-independence and open accessibility over earlier systems. As a result, there is ongoing research in the re-design and re-evaluation of AAS in order to transform them in task-oriented environments for design of algorithms in on-line mode. The modern task-oriented environments are classified as adaptive, intelligent and collaborative and implement different information and communication technologies. Users of task oriented environments in algorithms (TODEA) [5] have an opportunity to use a visual language for both subject and pedagogical knowledge description and generate programs with determined structure. Such an environment could provide students with teaching material, containing test-like exercises with tasks that could be performed by them. Each student could get not only a mark, but a complete list of errors, including conceptual mistakes committed while performing the task. An environment could be marked as adaptive if it is different for different individuals, teams, and groups by taking account information accumulated in individual, teams or group models. An intelligent environment applies the techniques from the field of Artificial Intelligence (AI) theory to provide broader and better support for the users. The aim of a collaborative environment is to support the activities of all individual participants and social groups in the teaching process [3]. User collaboration is a strict requirement for the teaching system to provide best possible results in achievement the pedagogic goal of the course.

A technology usually can be further dissected into finer-grain methods and techniques, which correspond to different variations of this functionality and different ways of its implementation [1]. The sequencing technology is implemented on high and low levels. High-level is ensured trough the opportunity to perform the exercise guided by the teachers' didactic skills and preferences. Low-level of this technology supports the learner

to perform a task construction in a way similar to those of the course author. Internal script languages on both levels ensure intelligent analysis of the learners' solution technology. Interactive task solving support technology also is implemented on two levels respectively on structural and functional knowledge. In such a way the learner and group's matching technologies will be added to those inherited from the WINDOWS-based prototype. According to the Brusilovsky & Peylo well known survey [1] the advanced e-learning systems known from the literature presents a combination of these features but only a few systems posses all these features.

The experimental study covers the different pedagogical and psychological aspects in computer science area. It allows the researchers to check the level of the acquired learning material and to evaluate how far didactic goals are satisfied. The result is quantitative dependence between different teaching parameters which are implemented in TODEA. [5]

The basic idea of this paper is to present the result of an experimental study, which exams the intelligent test and the conventional multiple choice test in the algorithm learning area. The main hypothesis is that the intelligent test covers more didactic goals, range more fully the taught material and leave the learners' better possibilities to show acquired skills in algorithms. The paper focuses on the following: goal, object and methodology of the study, experimental data, test quality evaluation, correlation analysis, and both test comparison.

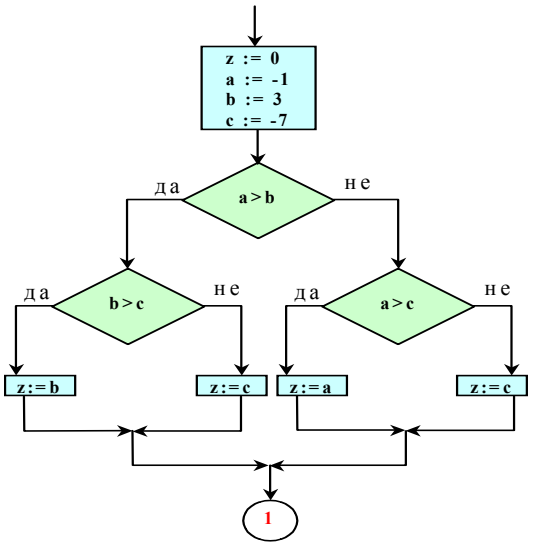
EXPERIMENT DESCRIPTION

The main goal of the study is to assess the effectiveness of the intelligent computer-based tests in the area of algorithms in comparison with the conventional ways of testing such as multiple-choice tests and exams. There were used techniques described in [6]. Object of the study were students-bachelor (1-st year, 1-st semester), specialty Communication Technique and Technologies at Rousse University. In the framework of the course Programming, part 1 the students learn some basic algorithms in computer science as block schemes and then they learn several topics on PASCAL programming algorithms The learning material on algorithms was presented in 2 lectures and applied in 8 practical exercises, where the students worked on computers in teams of 2-4. During the semester the students had to perform two written tests on the learned material and one course work concerning development the block scheme of an algorithm and its implementation in PASCAL. Then an integrated mark was formed on the base of the tests and coursework marks. If this mark was 5 or 6, and the student was satisfied with it he/she was relieved of the exam. Otherwise the student was examined during the session with a task similar to the course work.

Two tests covering the topic "Algorithms" were prepared in the form of WORD-documents. The first one - test (T) contains 25 multiple-choice questions and brings 25 scores respectively 1 score per the right answer (fig. 1). The second one – called intelligent (IT) contains 27 different types of questions with 250 total scores (fig 2). The used questions types are following: multiple choice, unordered keywords, ordered keywords, and unordered pairs [3]. It can be accepted that both tests have approximately one and same number of questions. At the same time IT has significant higher degree of covering the taught material.

The announcement for the T and IT had been made approximately a week before the experiment. 21 students-voluntaries appeared in a classroom to take a part in the experiment. They show personal interest and respect to become objective and precise knowledge assessment. They were introduced to the study goal, the way of answering the different questions and the way of their assessment.

ВЪПРОС 7 **:**
Каква ще бъде стойността на Z в т.1 от блок-схемата от фиг.1 ?
a) 3 **b)** 7 **c)** -1 **d)** 0
Отговор: …
Параметри: L=1; Q_t=1; C_p=1.
Препратка: 2.1

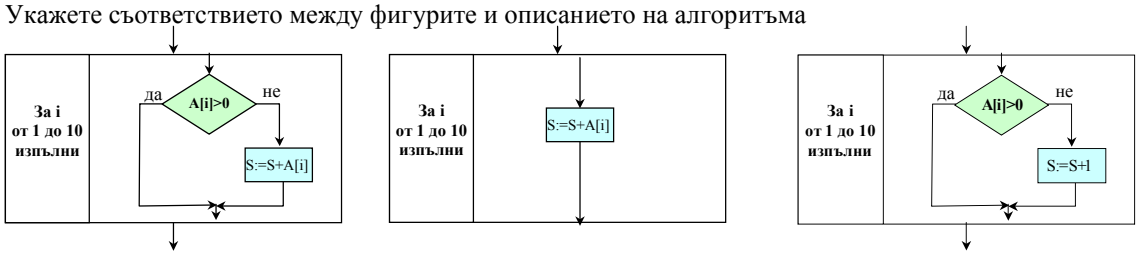


Фиг. 1

Fig 1. A question from the WORD document of the T

The main goal of the experiment was explained as a whole. The students were also told that the planned time is 20 minutes for the T, and 60 minutes for the IT. Actually the times for performance the tests were unlimited and the students have to register them.

ВЪПРОС 23: Неподредени двойки
Укажете съответствието между фигурите и описанието на алгоритъма



1) **2)** **3)**

а) брой на положителните елементи на масив А; **б)** сума от положителните елементи на масив А;
в) сума от елементите на масив А;
Отговор: .....
Параметри: L = 2; Qt=6; Cp=1.
Препратка: 5.

Fig 2. A question from the WORD document of the IT

DESCRIPTION OF THE EXPERIMENTAL DATA

The experimental data were brought to two tables (not shown here), containing the following information about each test – the student's name, the student's ID, the number of the points for each question, the total scores, the time, the mark and the final mark on

Prog.1. While the student's answers in the multiple-choice test T are 0 or 1 in intelligent one IT they are partially correct, e.g. in the range of [0,1]. The manually computed by the teacher students' tests scores and marks together with the students' time and mark on Prog. 1 are summarized in third table (Table 1). Following the methodology in [6] an additional processing the experimental results was made for test quality, group's knowledge evaluation, and both studies comparison.

Table 1. A part of the experimental data

I.1	Name	ID	Test	Scores	Time	Mark	Progr.1
1	Александър Харутюн Мурадян	083310	T	13	19	3	**4**
			IT	154	45	4	
2	Николай Цветанов Иванов	083301	T	16	17	4	**5**
			IT	201	64	5	
3	Петя Енчева Димитрова	083313	T	18	18	4	**5**
			IT	203	52	5	
....							
21	Гьонюл Бейхан Гайтан	083309	T	17	25	4	**5**

TEST QUALITY EVALUATION

According to the average questions scores of IT they were divided by difficulty into five groups respectively: very difficult (VD) – from 0 to 30, difficult (D) - 30-40, moderate (AV) - 40-60, easy (E) - 60-70, very easy (VE) - over 70. For each group the average result was calculated and the corresponding graphic was drawn using EXCEL. The resulted chart together with the average approximation is shown on fig.3 and fig. 4 respectively for T and IT. The main graphical characteristic of the question difficulty is the typical S-shape curve (fig. 5): 1) the width of the deviation by X-axis of a curve corresponds to the difficulty of the corresponding question, because too few correct answers are given; 2) the inclination describes the probability of guessing. The lower grade is, the more even is the curve of the distribution of answers. The average curve should be close the diagonal of the chart. It doesn't contain the whole information, that's why it's good to be combined with the number of questions in each group (fig.4). The first graphic is not suitable for T because all the answers have only 2 values: true or false. So, only the second graph type could do the juxtaposition for the both tests.

Shortly, the additional processing of the experimental data and their graphical interpretation confirm the expectation that both tests are valid

GRAPHIVAL DEPENDENCES

The circle graphical dependences of the students' marks for T and IT are shown on fig.6. It could be assert that most students achieved better result on the IT. Both tests were carried out one after one immediately and the students' knowledge was one and same. These facts demonstrate that the IT is better than T because its advantages, namely: more widely taught material covering and more aspects of knowledge assessment by means of the questions' parameters.

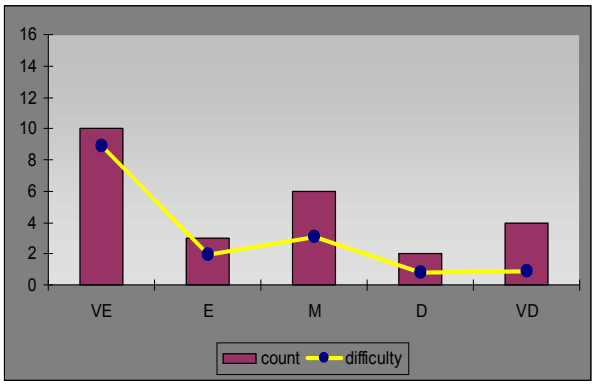


Fig.3. Question groups and their difficulty for T

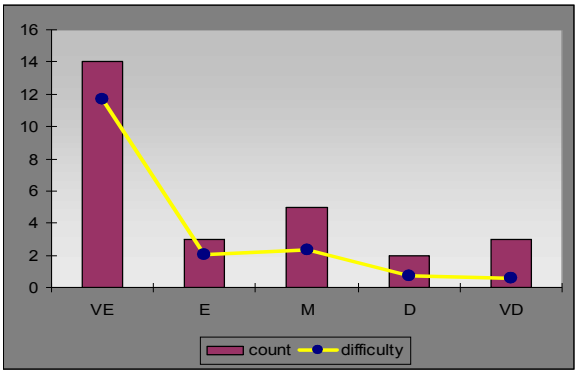


Fig.4. Question groups and their difficulty for IT

The experimental graphical interpretations of the dependences between the marks Prog1/T (fig.7) and Prog1/IT (fig.8) are presented with dots and their linear approximation with solid line. The smaller angle of the approximation line with respect to X-axis means lower sensitive regarding to the student's knowledge differences. From the visual comparing both figures follows the expected conclusion that the intelligent IT is more sensitive than multiple-choice T.

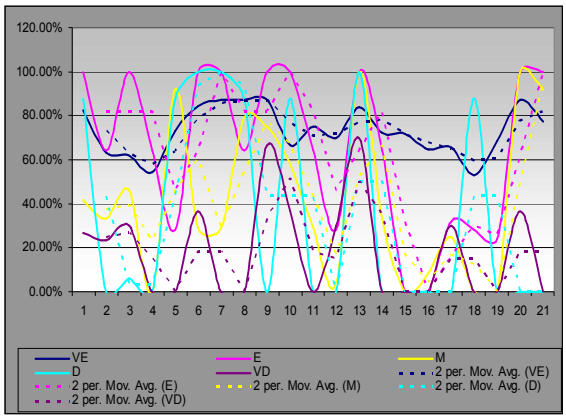


Fig.5. Questions' difficulty for IT

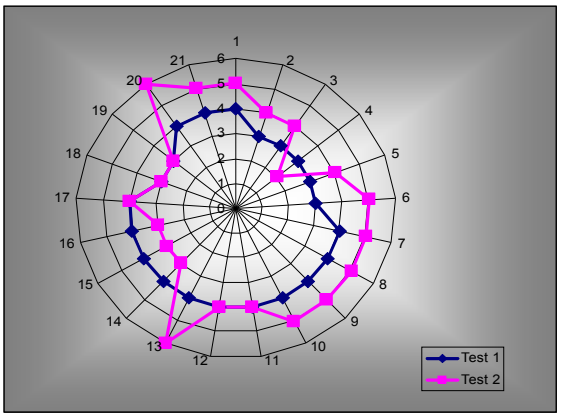


Fig.6. Students' marks for T and IT

The study also allows specifying the initial time for performance of both tests defines only by the authors' experience. The actual average time for T and IT was respectively 19 and 51 minutes. Because the time for T is very close to the planned one it could be increased to 25 minutes whereas the time for the IT can remain 60 minutes.

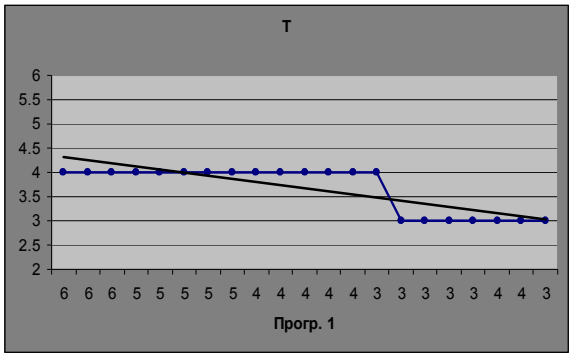


Fig.7. Mark's dependence Prog1 and T

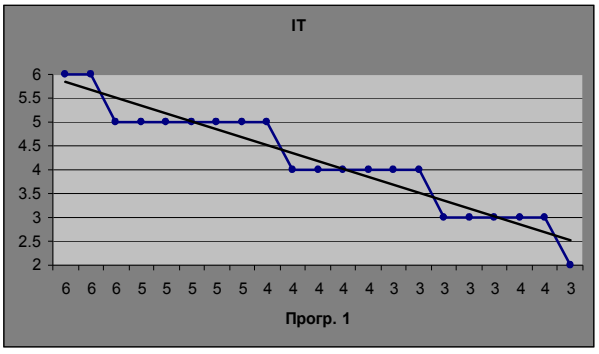


Fig.8. Mark's dependence Prog1 and IT

CORRELATION ANALYSIS

In general, the correlation analysis is used to measure the power of the relationship between two statistical factors. The linear correlation coefficient r, a real number in the range of [-1,1] can serve as a quality indicator for the relationship between both tests's parameters. For the current study the relationships between the following couples of parameters are of interest: Time T/MarkT, Mark T/Prog1, Time IT/Mark IT, Mark IT/Prog1 and Mark T/Mark IT. The corresponding correlation coefficients computed from the experimental data are given in table 2. The biggest correlation coefficient 0.786 for Mark IT/Prog1 means that there is high relationship between these two parameters. The relationship for Mark T/Prog1 and Mark T/MarkIT is moderate, and the one for Time T/MarkT and Time IT/Mark/IT is lower.

Table 2. Correlation coefficients for time and marks

Relationship	T: Time / Mark	Mark: T/Prog1	IT: Time / Mark	Mark: IT/Prog1	Mark: T/IT
Correlation coefficient	- 0.123	0.323	- 0.313	0.728	0.402

Table 3. Correlation coefficients for the two groups

Relationship	T: Time / Mark	Mark: T/Prog1	IT: Time / Mark	Mark: IT/Prog1	Mark: T/IT
Exellent students	0.164	-0.109	0.786	0.451	0.032
Mean level students	0.130	-0.063	0.201	0.247	0.191

According to the mark on Prog1 the experimental data were divided in two tables. The first one contains data for the excellent students (mark 5 and 6) and the second one for the mean level students (mark 3 and 4). That was made because the excellent students were assessed at the end of the first semester and released from exam. The results from the calculation of the same correlation dependences for the two groups of students are presented in table 3. Note that, the values of r close to 0 Time/Mark for mark T/Prog1 confirm the statement that is not objective due to some well-known reasons. For comparison r for IT/Prog1 has a positive value, greater for the excellent students than for the mean ones.

COMPARISON OF BOTH TESTS

Comparison of two different forms of sessions (in our case T and IT) on one and the same learning material can be useful for improving the questions formulation and updating the value of test's questions difficulty. To revise the IT questions and update the value of their degree of difficulty T and IT were compared (table 4). Significant difference between the values of difficulty for 3 questions (dark grey) could be explained with unclear formulation of the question, possible alternatives for keywords, and other reasons. So these questions have to be revised or deleted from the question base. For 10 questions (light grey) this difference is insignificant and for the rest ones there is no difference (not highlighted). The values of the test questions difficulty in the question base have to be updated. The total test difficulty 0,59 means that the IT has moderate degree of difficulty. That is also confirmed by the corresponding correlation coefficient r=0,786. It is expected that each new experiment with IT will increase this value.

Table 4. Questions degree of difficulty for both kind of tests

Experiment	Q1	Q2	Q3	Q4	Q5	Q6	Q7	Q8	Q9	Q10	Q11	Q12	Q13
T	0,90	0,24	0,80	0,52	0,62	0,57	0,90	0,80	0,80	0,90	0,80	0,90	0,80
IT	0,95	0,81	0,10	0,70	0,76	1,00	0,71	0,80	0,20	0,80	0,80	0,90	0,80
Updated	0,90	0,50	0,45	0,60	0,70	0,75	0,80	0,80	0,50	0,85	0,80	0,90	0,80

Experiment	Q14	Q15	Q16	Q17	Q18	Q19	Q20	Q22	Q23	Q24	Q25	Q26	Q27
T	0,70	0,70	0,80	0,60	0,30	0,00	0,20	0,40	0,80	0,70	0,30	0,40	0,40
IT	0,80	0,70	0,80	0,40	0,50	0,30	0,30	0,40	0,70	0,90	0,30	0,60	0,50
Updated	0,75	0,70	0,80	0,50	0,40	0,15	0,25	0,40	0,75	0,80	0,30	0,50	0,45
												Total test difficulty:	**0,59**

Observation of the students' normalized marks for T and IT (table 5) provides information about the level of didactic goals which both test's question cover. It might see that the results from IT are better for the most students. The main reason for this is that questions of the IT give some additional opportunities for learner and teacher. Learner can show more fully his knowledge and abilities. The teacher can check the level of different didactic goal achievement be means of question parameters, namely: L – level of the complexity, Cp – degree of prompt and D – degree of difficulty. There are two students - St4 and St15 which had shown better results on T. The main reason is due to T's chance to answer by good fortune. The IT avoids this fault. Its answers have fill in requirement and the learner has to show knowledge and creativity.

Table 5. Normalized marks for T and IT

	St1	St2	St3	St4	St5	St6	St7	St8	St9	St10	St11
T	0.68	0.52	0.40	0.52	0.52	0.52	0.64	0.72	0.76	0.64	0.64
IT	0.82	0.60	0.65	0.38	0.62	0.84	0.80	0.81	0.86	0.75	0.63

	St12	St13	St14	St15	St16	St17	St18	St19	St20	St21	Ave
T	0.64	0.76	0.72	0.68	0.64	0.60	0.48	0.56	0.72	0.68	0.62
IT	0.58	0.94	0.53	0.42	0.48	0.58	0.49	0.51	0.86	0.74	0.66

CONCLUSIONS

As in the case of WEB technology the difference between the expected and real results from both multiple-choice and intelligent tests is insignificant, e.g. both tests are valid. However the usage of WORD technology increases the teacher's labor cost for computing instead the environment. In comparison with the multiple-choice test the intelligent one is more effective regarding the level of the tested knowledge, degree of covering the taught material, length of the test, and sensibility of the learner' s knowledge assessment. The relationship between the final mark given by the teacher and the intelligent test mark is very close, while the relationships between the differential marks and multi-choice test vary from moderate to low. The reported results confirm the hypothesis that intelligent test in algorithms gives up more abilities for learner and teacher in their actions and avoids faults of the conventional multiple choice T. A tool for post processing the learner's test results will be implemented as a part of an intelligent and adaptive task oriented teaching environment in algorithms.

REFERENCES

[1] Brusilovsky, P., (1999), Adaptive and Intelligent Technologies for Web-based Education, Special Issue on Intelligent Systems and Teleteaching, Kunstliche Intelligenz, Vol. 4, pp. 19-25.

[2] Marcelino, M., A. Gomes, N. Dimitrov, A. Mendes. Using a Computer-based interactive system for the development of basic algorithmic and programming skills. International Conference on Computer Systems and Technologies (e-Learning), Rousse, Bulgaria, 2004, 17-18 June, IV.8-1 – IV.8-6.

[3] Zheliazkova I.I., Andreeva M.H., An Intelligent Multimedia Environment for Knowledge Testing, E-learning and the Knowledge Society, Gent & Brussels, 2004, pp.3.13.1-3.13.24.

[4] Zheliazkova I, Atanasova G., Computer-Aided Teaching and Learning Algorithms (A survey). Proceedings of the European Conference "Innovation in Education for Electrical and Information Engineering", Sofia, 2004, pp. 49-58.

[5] Zheliazkova I.I., Atanasova G.E., A Visual Language for Algorithm Knowledge, Proceedings of the International Conference on Computer Systems and Technologies (e-Learning), Rousse, Bulgaria, 17-18 June, 2004, pp.IV.24-1-IV.24-6.

[6] Zheliazkova I.I., Kolev R.T., Multiparametrical Analysis of Test Results, International Conference on Computer Science'05, Chalkidiki, Greece, 30 September- 2 November, 2005.

CONTACT ADDRESS

Sen. Assist. Galina Atanasova
Department of Informatics and Information Technologies
Angel Kanchev University of Ruse
Phone: (+359 82) 888 326
E-mail: gea@ami.uni-ruse.bg

ТРАДИЦОННАТА ТЕСТОВА ПРОВЕРКА НА ЗНАНИЯ В СРАВНЕНИЕ С ИНТЕЛИГЕНТНИЯ ТЕСТ В ОБЛАСТТА НА АЛГОРИТМИТЕ – ЕКСПЕРИМЕНТАЛНО ИЗСЛЕДВАНЕ

Галина Атанасова

Русенски университет "Ангел Кънчев"

Резюме: *Основната идея на статията е да представи резултатите от експериментално изследване в областта на алгоритмите. Сравнени са два начина за проверка на знанията в тази област: традиционния тест с многовариантен избор и интелигентен тест. Доказана е основната хипотеза, че интелигентния тест в областта на алгоритмите покрива по-пълно учебния материал, дава възможност за реализиране на различни дидактически цели и предоставя възможности на обучаваните да покажат по-добре придобитите възможности от обучението си.*

Ключови думи: *компютърни науки, алгоритми, задачно-ориентирани среди, интелигентно тестване, експериментално изследване.*

ISSN 1311-9184