



Lodz University of Technology, Lodz, Poland
Faculty of Electrical, Electronic, Computer and Control Engineering

A Web System for Assessment of Students' Knowledge

**Lidia Jackowska-Strumiłło, Wojciech Bieniecki
and Marwah Bani Saad**

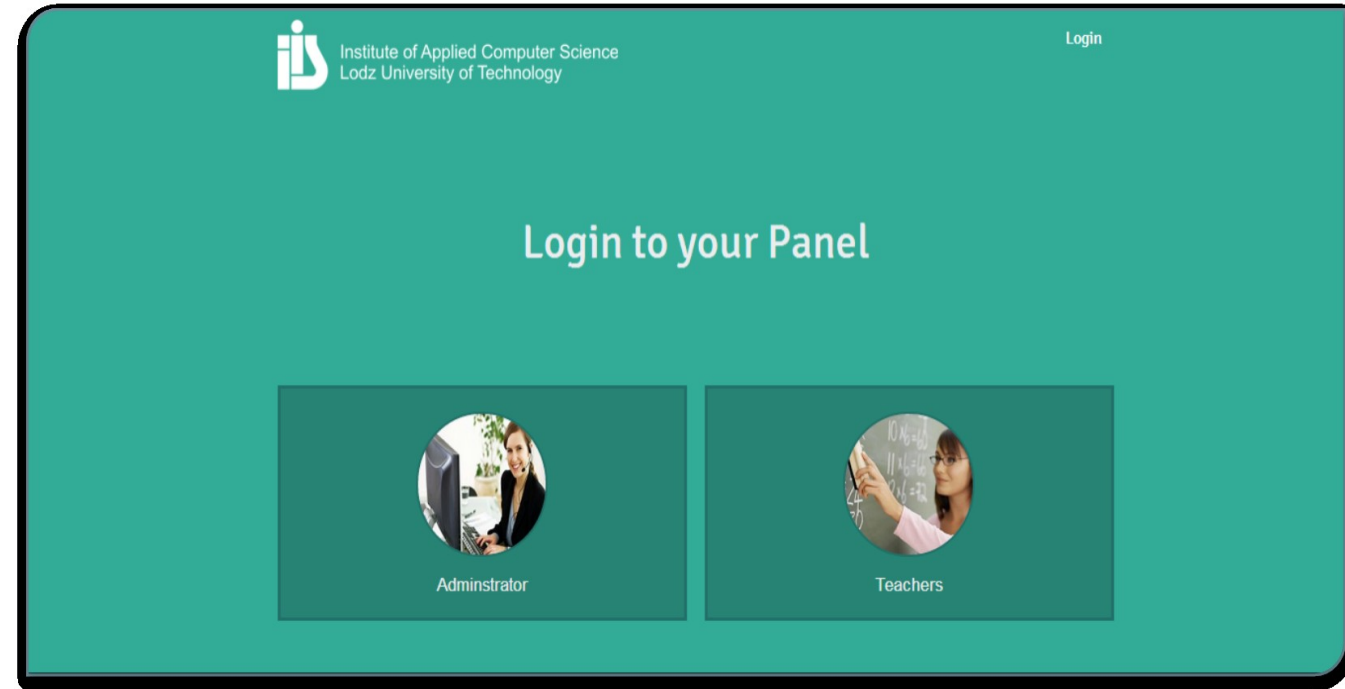
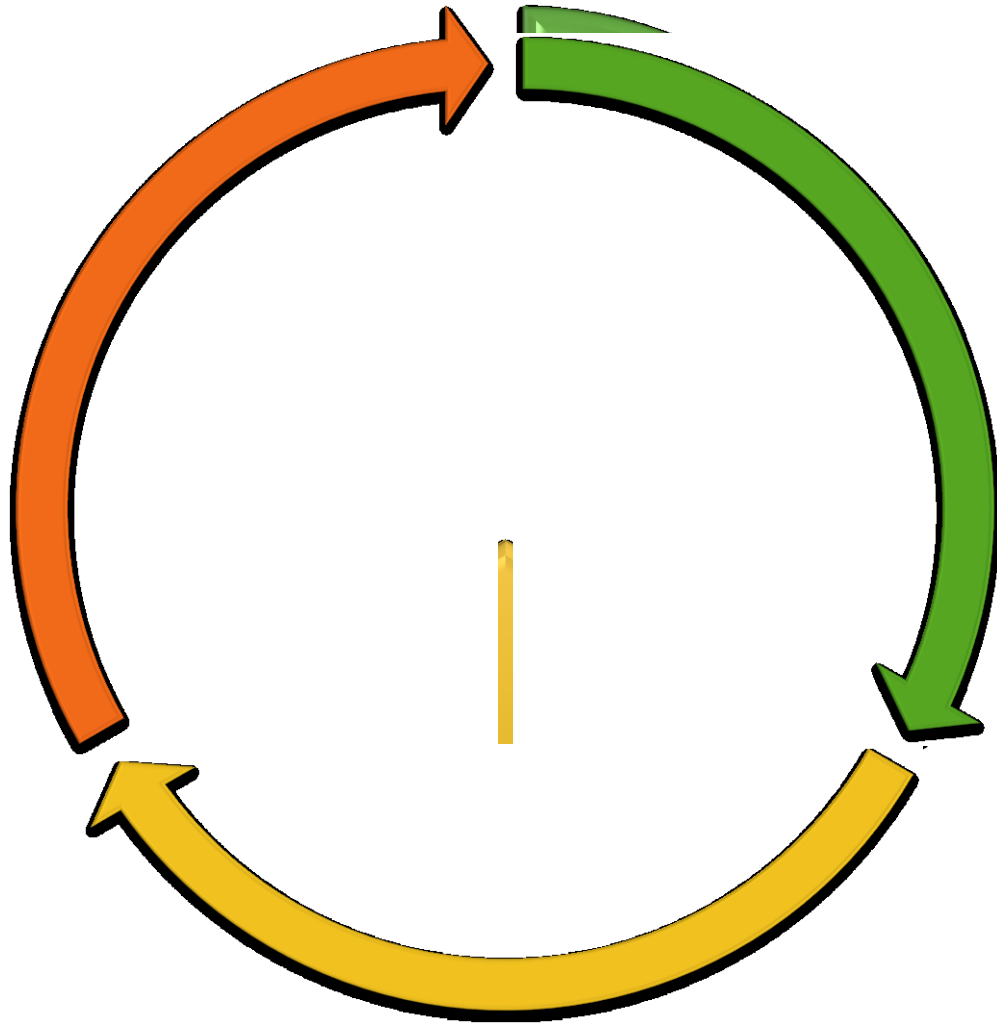


Institute of Applied Computer Science

Aims

- ▶ E-test system building and testing
- ▶ Research problem
 - ▶ Adding a new type of open questions
 - ▶ Creating a new type of scoring
 - ▶ Implementing new text processing algorithms

E - Test General Overview



The construction of the tests

Creation
of
categories

Creation
of
questions

Test
parameters
setting

Test planning

Category	No.Q	No.Q/Cat
C2_fun	8	<input type="text" value="2"/>
C2_rek	8	<input type="text" value="2"/>
C2_tablice	9	<input type="text" value="1"/>
C2_org_zas_zm	7	<input type="text" value="1"/>
C2_instr_sterujace	7	<input type="text" value="1"/>
C2_fun_bibl	9	<input type="text" value="1"/>
C2_instr_sterujace2	5	<input type="text" value="1"/>
algorytmy2	7	<input type="text" value="1"/>
All Questions : 60		Randomized Questions : 10

Percentage Grade Scale

ndst dst dst1/2 db db1/2 bdb

pts: 5 0 8 0 10

Ratings half Quiz

Password	Data	Valid from	Valid until	Edit	Remove
ala	25 01 2015	17:30	23:50		
Gt56vz	26 01 2015	12:15	12:45		
DF89u6g	27 01 2015	14:15	14:45		
Fs34nb7	28 01 2015	14:15	14:45		
Mj72xp	29 01 2015	10:15	10:45		
Hc63kd	29 01 2015	12:15	12:45		
Juo587	30 01 2015	10:15	10:45		

/ / : :

March, 2015

Mon	Tue	Wed	Thu	Fri	Sat	Sun
						1
2	3	4	5	6	7	8
9	10	11	12	13	14	15
16	17	18	19	20	21	22
23	24	25	26	27	28	29
30	31					

April, 2015

Mon	Tue	Wed	Thu	Fri	Sat	Sun
		1	2	3	4	5
6	7	8	9	10	11	12
13	14	15	16	17	18	19
20	21	22	23	24	25	26
27	28	29	30			

Times

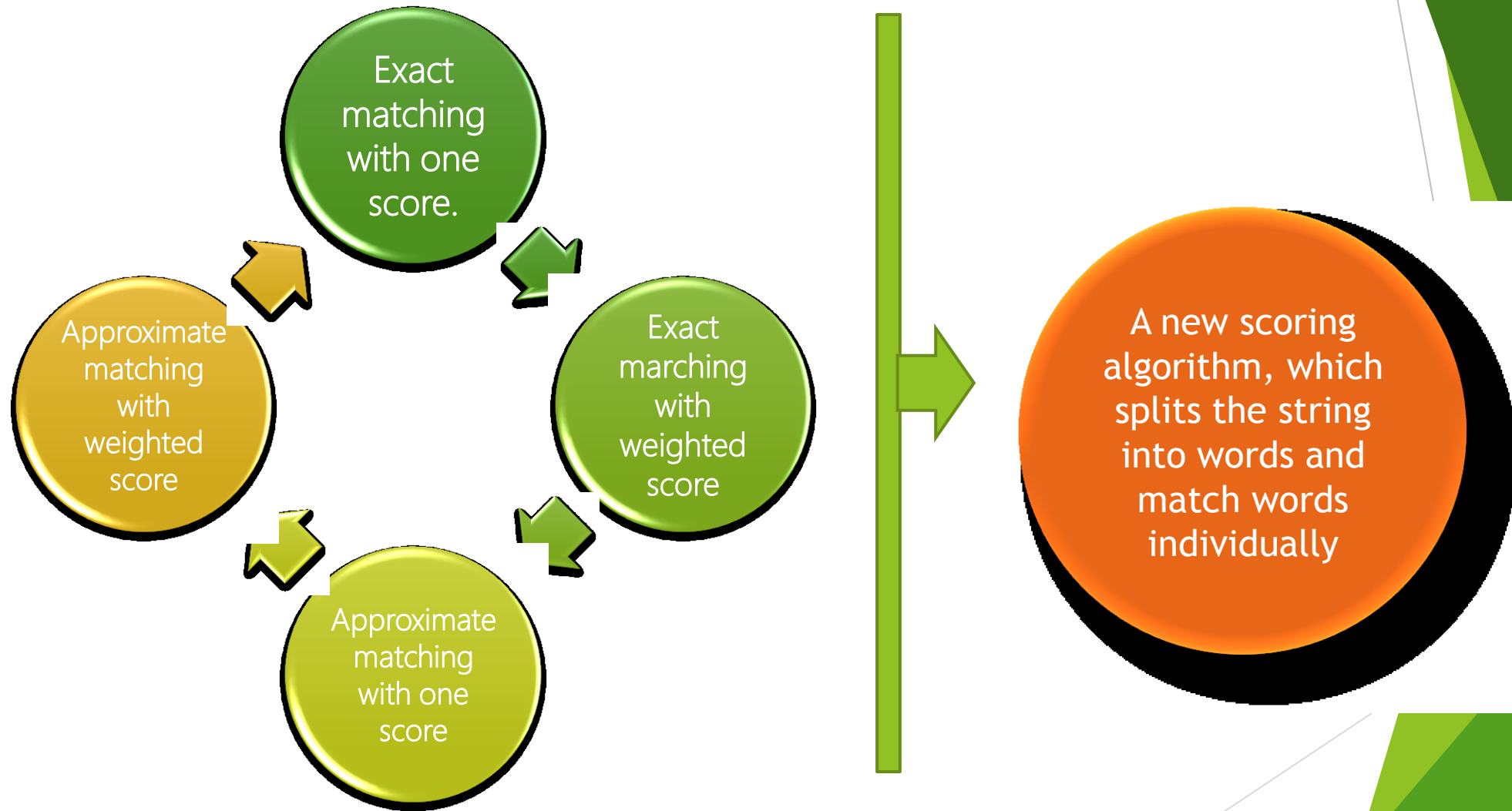
Time to reply. [sec.]:

Time to test [min.]:

Question Types

1. Multiple choice question (one of four)
2. Short-Answer Question
 - ▶ Question: **What is blue, or red, or yellow?**
 - ▶ Acceptable Answers:
 - ▶ a colour
 - ▶ a color
 - ▶ it's a colour
 - ▶ it's a color

New Scoring Process



Exact matching + one score.

- ▶ **The question:** *One KiB contains ...*
- ▶ **The four possible answers** (not displayed to the student):
 1. *1024 B*
 2. *1024 bytes*
 3. *1024 byte*
 4. *1024B*
- ▶ **The student's answer:**

1 KB = 2¹⁰ B = 1024 B (bytes) (Fault F1)
- ▶ **Algorithm of Exact Comparison** using (= =)

Exact marching + weighted score

▶ **Question: What is the minimum passing score?**

▶ **Acceptable Answers:**

1. 3
2. 3.5
3. 4
4. 5

▶ **Possible Answers:**

- | | | |
|----|-----|-------------|
| 1. | 3 | Score= 100% |
| 2. | 3.5 | Score= 80% |
| 3. | 4 | Score= 60% |
| 4. | 5 | Score= 40% |

▶ **Algorithm of Exact Comparison**

Approximate matching + weighted score

▶ **Question:** What is a rabbit?

▶ **Acceptable Answers :**

1. Animal
2. Mammal
3. Vertebrate

▶ **Possible Answers:**

Vertebrate	Score=100%
Anemal	Score=70%
Fish	Score= 0%

▶ **Algorithms for Approximate Equality (Levenshtein)**

The Levenshtein algorithm

$$k = \arg \min_{i=1..4} (\text{Levenshtein}(s, \text{acc_ans}_i), i)$$

$$l = \max(\text{len}(s), \text{len}(\text{acc_ans}_k))$$

$$c\% = \frac{l - \text{lev}(s, \text{acc_ans}_i)}{l}$$

where: *s* - filtered student's response

acc_ans - an array of up to four acceptable answers

l - the larger of lengths of strings (*s* and *acc_ans_k*)

c - the calculated compliance ratio

Approximate matching + one score

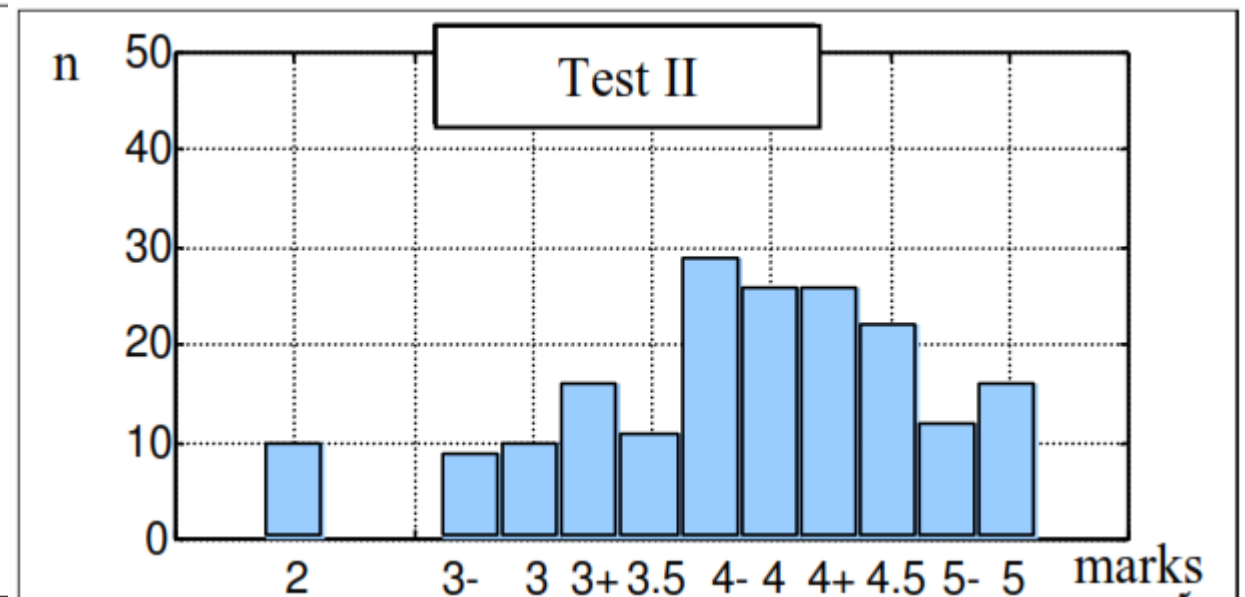
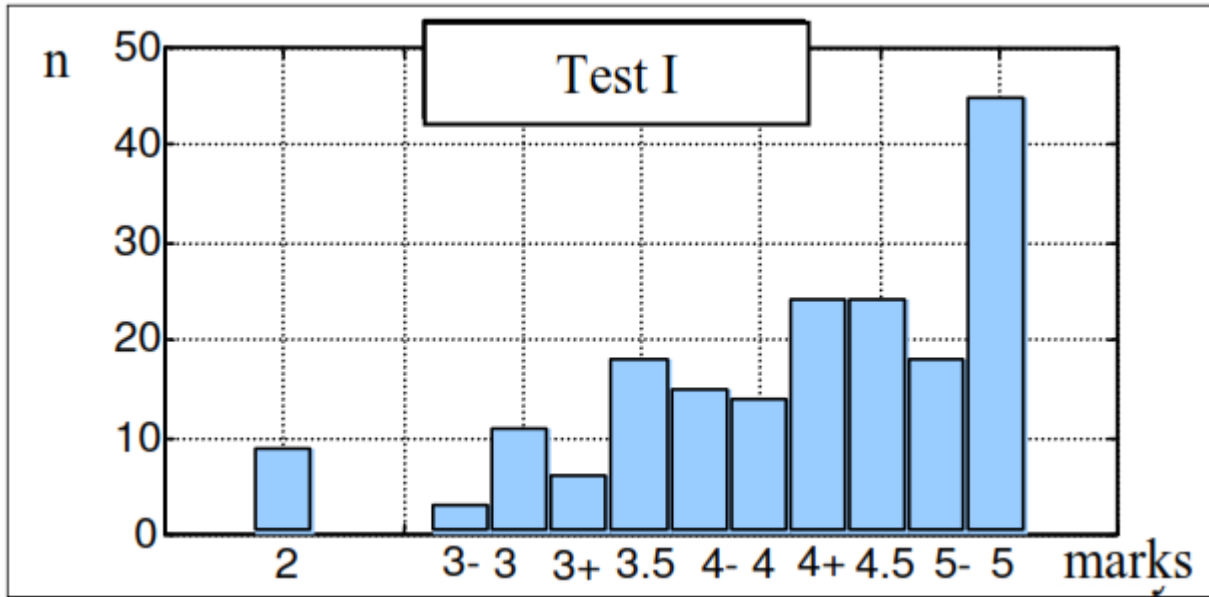
- ▶ **The question:** *DMA transmission is controlled by:*
- ▶ **The four possible answers** (not displayed to the student):
 1. *DMA controller*
 2. *controller of DMA transmission*
 3. *DMA transmission module*
 4. *DMA module*
- ▶ **The student's answer:**
interrupts controller (Fault F2)
- ▶ **Algorithms for Approximate Equality** (Levenshtein)

Algorithm of scoring for longer answers

```
score ← 0
st_ans ← GetEditArea( )
s ← Filter (st_ans)
FOR EACH ans IN acc_ans
    points ← 0
    array ← Split(ans)
    FOR EACH word IN array
        IF s CONTAINS word
            points ← points +1
        END IF
    END FOR
    points ← points / Size(array)
    IF points > score
        score ← points
    END IF
END FOR
```

System Testing 1

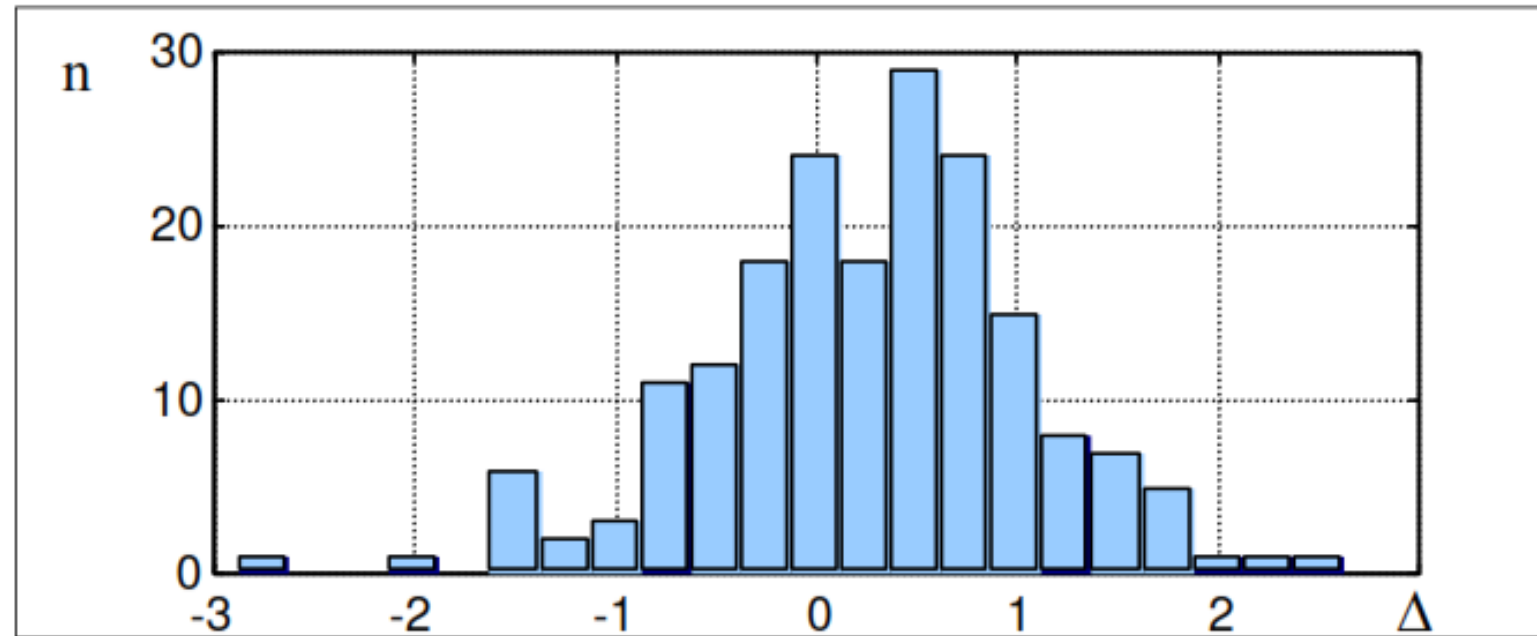
close and open questions



Histograms of the marks of tests I (open questions) and II (close questions)

System Testing 1

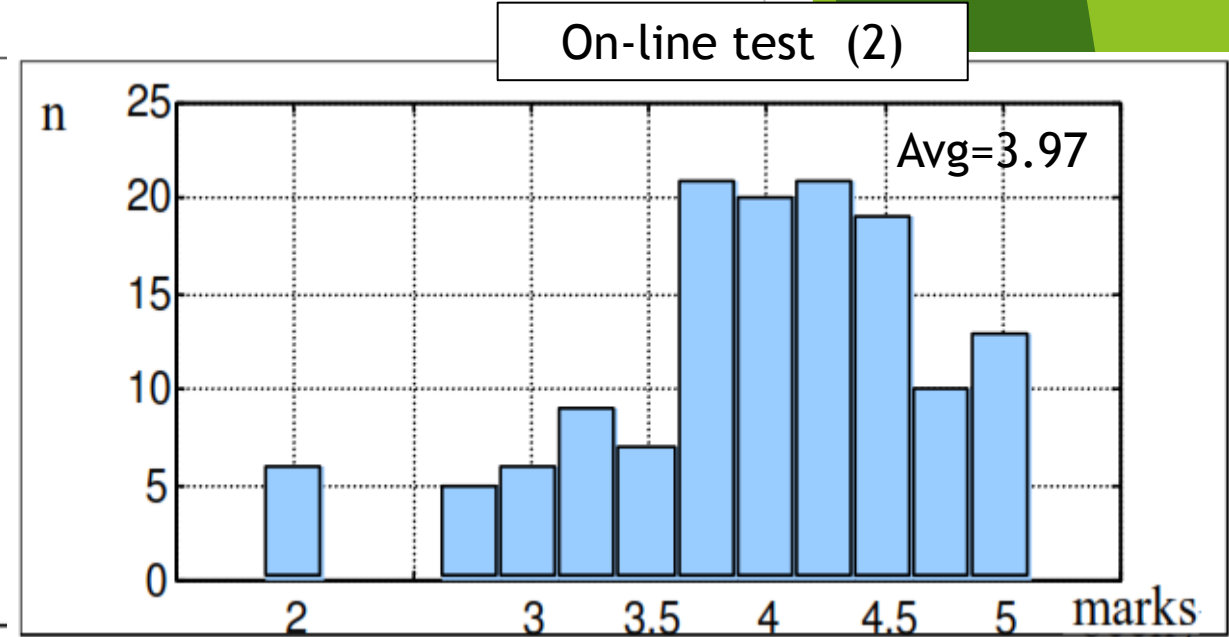
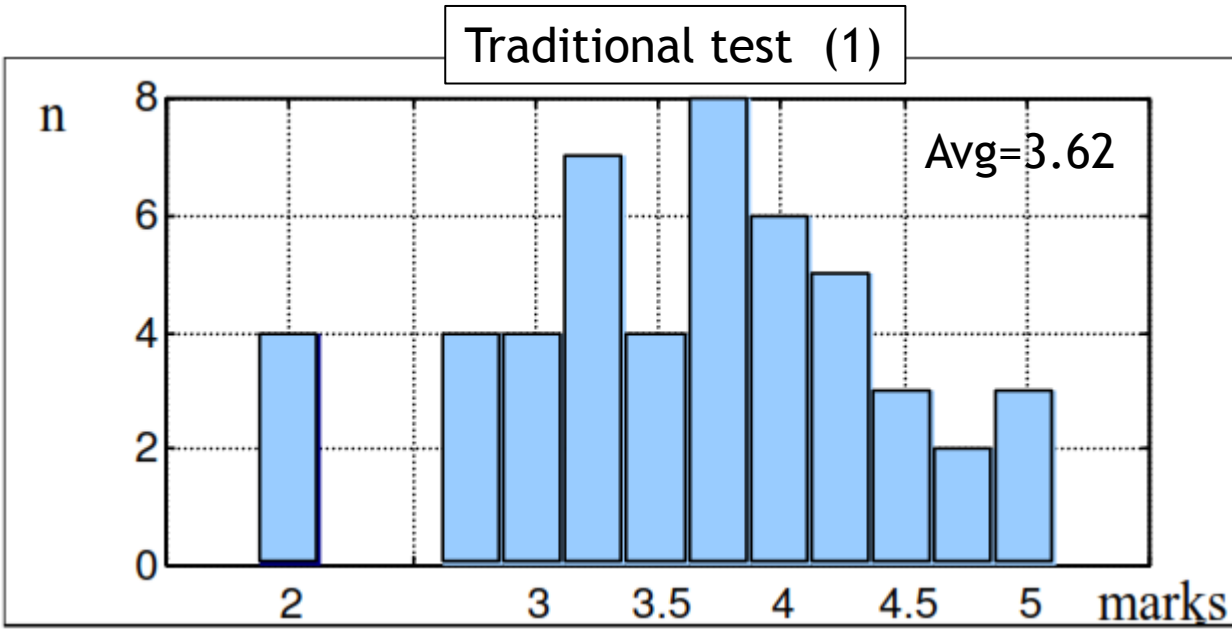
close and open questions



Histogram of differences between scores of the first and second test

System Testing 2

traditional and on-line test



Histograms of marks for two groups of students who wrote the second test on the paper (1) or on the computer (2)

System Testing 3 on-line open questions

TABLE I. SHORT ANSWER QUESTIONS WITH EXACT MATCHING

Number of questions	Type of questions	Verification of scoring algorithm		
		Correct score	Fault F1	Fault F2
67	T1	58	9	0

Fault F1 - correct answers are not recognized; fault F2 - wrong answers are recognized

TABLE II. SHORT ANSWER QUESTIONS WITH APPROXIMATE MATCHING BY THE USE OF THE LEVENSHTAIN ALGORITHM

Number of questions	Type of questions	Verification of scoring algorithm		
		Correct score	Fault F1	Fault F2
397	T2	371	19	7

Fault F1 - correct answers are not recognized; fault F2 - wrong answers are recognized

System Testing 3

new algorithm

TABLE III. SHORT ANSWER QUESTIONS WITH INDIVIDUAL WORDS MATCHING BY THE LEVENSHTAIN ALGORITHM

Number of questions	Type of questions	Verification of scoring algorithm		
		<i>Correct score</i>	<i>Fault F1</i>	<i>Fault F2</i>
397	T2	389	8	0
67	T1	66	0	1

Fault F1 - correct answers are not recognized; fault F2 - wrong answers are recognized

Split-approximate matching + one score

▶ **The question:** *One sector in the floppy disc has a size of...*

▶ **The four possible answers** (not displayed to the student):

1. *512 B*

2. *512 bytes*

3. *512 byte*

4. *0.5 KiB*

▶ **The student's answer:**

512kB (Fault F2)

▶ **Algorithm with input string splitting and Approximate Equality (Levenshtein)**

Conclusion

- ▶ E-test system was built and tested
- ▶ A new concept of applying short answer questions introduced
- ▶ For the exact matching scores 14% correct answers were not recognized.
- ▶ The Levenshtein measure is inadequate for comparing answers consisting of a few words (7% faults)
- ▶ A new scoring algorithm, which splits the string into words and matches words individually, has reduced the faults to 2%