INTERNATIONAL COMPARISON OF PROBLEMS FROM AN INFORMATICS CONTEST

Peter Tomcsányi and Jiří Vaníček

Abstract: Both Slovakia and Czech Republic joined the international informatics contest Bebras (Beaver in English). The contest was conducted under the respective names *iBobor* and *Bobřík informatiky*. This gave us the opportunity to select the problems, which were identical or very similar in both countries and compare the achievements of the participants as well as their ways of solving the problems.
Our attention has been drawn to the fact that some of the problems had the same results in both countries while others have shown significant differences. Therefore we tried to find and explain the reasons of these differences. Sometimes it is a tiny but significant change in the text of the problem to be solved, sometimes the problems were given to slightly different age groups, sometimes the only difference is the number of participants, which decided not to try solving the problem (not to lose points when giving an incorrect solution). The results of our comparison can be used both as impulses for preparing problems for the next runs of the contest and inform the broader scientific community.

Key words: Informatics contest, primary and secondary schools.

1 INTRODUCTION

The national implementations of the international Bebras contest in informatics [1] in Czech Republic and Slovakia in the school years 2007/8 and 2008/9 used the same sources of problems for the contest - Bebras international databases of problems.

But the contests in Czech Republic and Slovakia were not 100% identical. When comparing the problems from two runs of the contest in Slovakia and one run in the Czech Republic we found 10 problems, which were identical or very similar. In this paper we try to compare their results and try to explain the observed differences in results.

Before doing so, we would like to stress that the primary purpose of the contests was not collecting data for exact scientific comparisons. For example, the local organizers did not pay special attention to using exactly the same wording in both countries. Also the categories of participants are similar, but not identical. Moreover, in Slovakia the categories were changed between the first and the second year. Also the numbers of participants differ, but there were several hundreds of participants in each category in each country. Therefore our comparisons sometimes cannot give exact quantitative answers. Instead they give questions and ideas for the next years of the contest.

To test statistical relevance of differences, we used sign schemes in contingency tables based on z-scores using the procedures and formulae described in [2].

There were 2 problems, which were solved in both countries with no differences in the number of correct and incorrect solutions. Another problem, named Park the Car, required the contestant to give instructions to a car (a sequence of *left*, *right* and *forward* commands) in order to navigate it into the parking place. The results are somewhat (statistically significantly) better in the Czech republic and we do not see any special reason for that.

In the rest of the problems we observed several kinds of differences, which brought interesting ideas described in next sections.

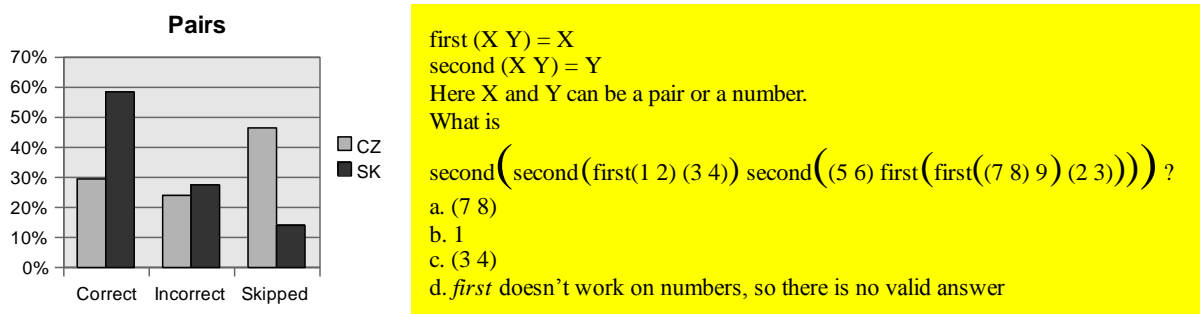## 2 "SKIP IF NOT SURE" SOMETIMES WORKS, SOMETIMES DOESN'T

The scores achieved in several problems suggest that Czech participants tended towards not answering the question if they were not sure with the answer more often than the Slovak participants did. This approach is rational because for an incorrect answer there is a penalty while for not giving an answer there is no penalty. The easy but somewhat cheap explanation would be that Slovakian participants are more competitive. But we also know that in Czech republic there has been a bigger emphasis on explaining the usefulness to skip a problem if the participant is not sure with the answer.

The interesting point about the strategy "skip if not sure" is that it seems that sometimes it has brought a better result, sometimes not.

Typical example of the first case can be seen in the Ciphering problem for Benjamins and in the Set the Table problem for Juniors. The difference in the number of successful participants is statistically irrelevant while the difference in the unsuccessful participants is clearly relevant. From the following graph we can see that the extra number of skipped answers in Czech republic roughly equals to the "missing" incorrect answers:
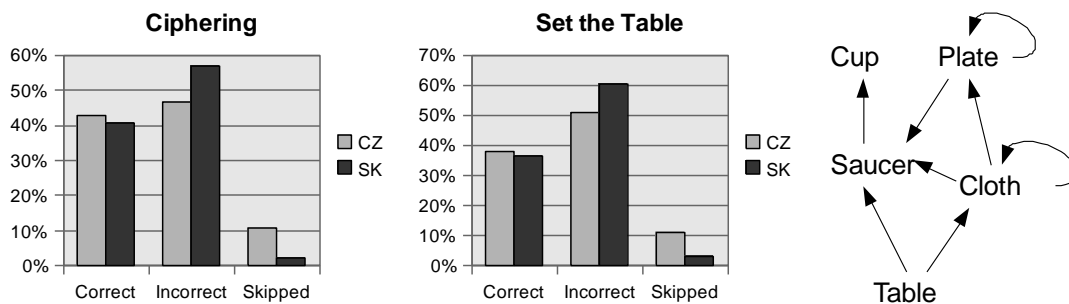
The other case can be seen in three problems, the effect is the strongest in the Seniors problem named Pairs. This time the differences in the number of incorrect answers are statistically irrelevant and the extra number of skipped answers in the Czech republic is "missing" from correct answers:

first $(X\ Y) = X$
second $(X\ Y) = Y$
Here X and Y can be a pair or a number.
What is

$$\text{second}\Big(\text{second}\big(\text{first}(1\ 2)\ (3\ 4)\big)\ \text{second}\big((5\ 6)\ \text{first}\big(\text{first}((7\ 8)\ 9)\ (2\ 3)\big)\big)\Big)\ ?$$

a. (7 8)
b. 1
c. (3 4)
d. *first* doesn't work on numbers, so there is no valid answer

Many Czech participants were deterred from trying to solve the problem even if the text of the problem included coloured parentheses (or was it exactly the colourful text, which suggested to them that the problem is too complicated to solve?). The Slovak text only included parentheses of different height like the original text of the problem shown above.

The same effect, but with lower intensity, can be seen in another two problems: Computer Description and Logical Operations on Pictures.

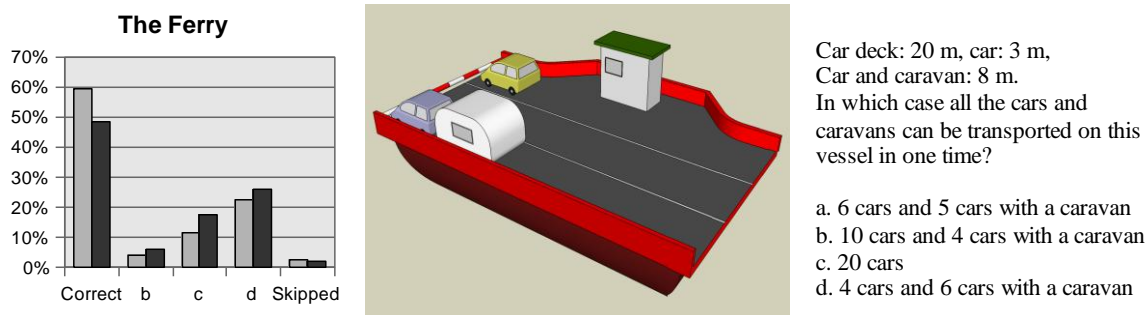## 3 WHAT CAN HAPPEN IF THE TEXT IS SLIGHTLY DIFFERENT

In two problems we can observe the interesting effects of differences between Slovak and Czech text.

First such problem is the already mentioned Set the Table. The difference here were the pictures of the tables. The contestant had to choose the one that was set correctly according to the rules (see the picture of rules in the previous section). While the Czech text included the original pictures from the international database, the Slovak team decided to re-paint the pictures both because the colour of the table was too similar to the colour of the cloth and because it was not easy to see that the picture of answer b shows a cup put inside another cup.

Set the Table

The Czech answer b   The Slovak answer b
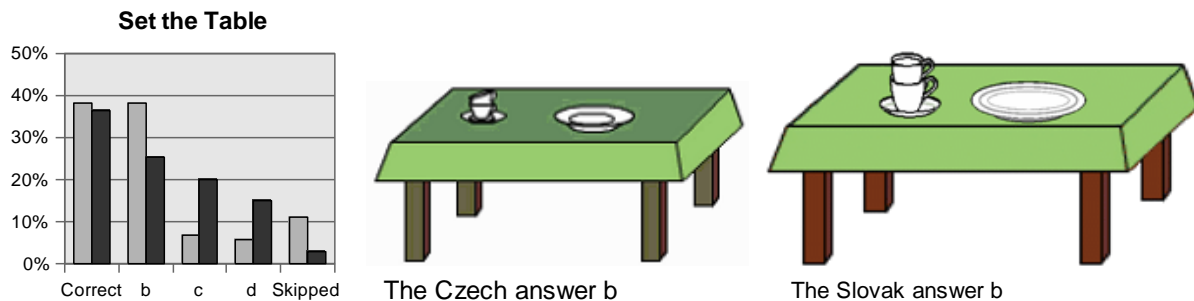
The change seems not to affect the number of correct answers - they are not statistically significantly different. It had some effect on the number of skipped answers (as we have seen above in the section 2 of this paper). But the change significantly affected the frequencies of incorrect answers. Much more Czech participants have preferred answer b, while in Slovakia it was easier to see that answer b is incorrect due to the changed picture.

In The Ferry problem, the change in the text of the problem had a different effect. The original text of the problem taken from the international database asked how many cars and cars with caravans can fit onto a ferry. In the Netherlands (the originating country of the problem) both ferries and caravans are more usual than in Slovakia or Czech republic. But the main problem was, in our opinion, with the wording - "cars and cars with caravans" repeated several times in the translated text. This somewhat clumsy phrase can be easily misinterpreted that the car can be detached from the caravan in order to fit on the ferry.

The Ferry

Car deck: 20 m, car: 3 m,
Car and caravan: 8 m.
In which case all the cars and caravans can be transported on this vessel in one time?

a. 6 cars and 5 cars with a caravan
b. 10 cars and 4 cars with a caravan
c. 20 cars
d. 4 cars and 6 cars with a caravan

In Slovakia the organizers left it as in the original English version, while the Czech organizers decided to change cars with caravans to buses both getting a shorter phrase "cars and buses" and avoiding thinking about their detaching. They also omitted the picture, which accompanied the original problem. The results suggest that these changes may contributed to the better results of Czech participants while the change seem to have no significant effect to to the relative frequencies of the incorrect answers.

4 THE AGE MATTERS

The last problem in our comparison was about choosing a right subject line for an e-mail message. In both countries it has been given to Benjamins, but the age interval for Benjamins in Slovakia in the year 2008/9 has been quite different than in the Czech republic. Thanks to this fact it is hard to compare the results, but this problem can be used for comparing how the relatively small differences in age can change the overall picture of a problem's difficulty.

The first picture shows the complete results. The second picture shows the results when Czech 8th graders are not considered and the third picture shows the results if only 5th and 6-graders and the corresponding lower secondary grades are considered.
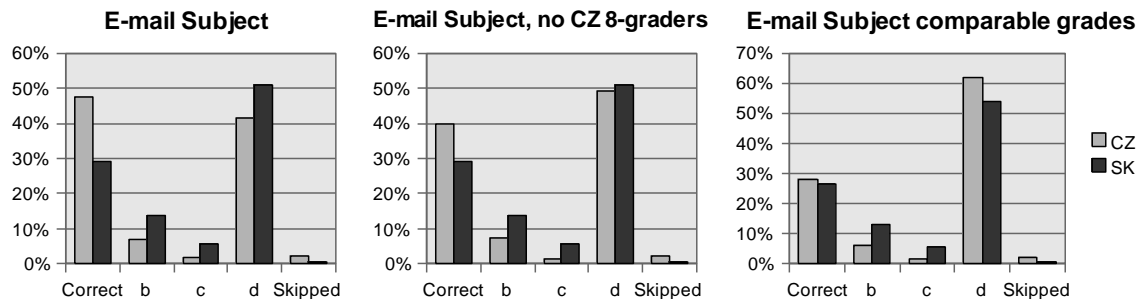
We can see that while the first picture shows a big difference in favour of Czech contestants, the second picture shows somewhat smaller difference and the last picture, which actually compares only students of the same classes and ages, shows no statistically significant differences between correct and incorrect answers. We can also see that with decreasing frequency of correct answers the frequency of answer d is increased. Answer d contained the whole message. The Czech version even contained "Thanks" at its end, so maybe therefore the Czech contestants liked this answer even somewhat more than Slovakian contestants.

We draw the conclusion that most of the younger contestants do not know what is a subject line of an e-mail. So the problem was not suitable for the 5th and 6th graders. And, more importantly, we must take big care when comparing the results especially in the younger age groups where one year of difference in age can make big difference in knowledge.

5 CONCLUSIONS

We have presented some interesting facts, which we discovered by comparing the results of international Bebras competitions in Czech Republic and Slovakia. We have seen that while some problems were solved with the same success in both countries, some others were not. For all such differences except one we tried provided an explanation or a hypothesis.

The general conclusion is that for the future we should try to make such comparisons easier by unifying the texts of some number of problems and also by adding some problems into the next runs, which would be constructed specifically for studying a well defined hypothesis.

The research is not finished, it will go on in the next years of our national Bebras contests and it will also drive the process of creating problems for the next years.

REFERENCES

[1] Dagienė V.: Competition in Information Technology: an Informal Learning, In:Proceedings Eurologo 2005, Warsaw, Poland pp. 228 - 234

[2] Chráska, M.: Metody pedagogického výzkumu, Grada, Praha 2008, ISBN 978-80-247-1369-4

Peter Tomcsányi
Univerzita Komenského, Fakulta matematiky, fyziky a informatiky
Mlynská dolina, 842 48 Bratislava
tomcsanyi@slovanet.sk

Jiří Vaníček
Jihočeská univerzita v Českých Budějovicích, Pedagogická fakulta
Jeronýmova 10, České Budějovice
vanicek@pf.jcu.cz