

Score Calculation in Informatics Contests Using Multiple Criteria Decision Methods

Jūratė SKŪPIENĖ

Vilnius University Institute of Mathematics and Informatics
Akademijos 4, LT-08663 Vilnius, Lithuania
e-mail: jurate.skupiene@mii.vu.lt

Received: August 2010

Abstract. The Lithuanian Informatics Olympiad is a problem solving contest for high school students. The work of each contestant is evaluated in terms of several criteria, where each criterion is measured according to its own scale (but the same scale for each contestant). Several jury members are involved in the evaluation. This paper analyses the problem how to calculate the aggregated score for whole submission in the above mentioned situation. The chosen methodology for solving this problem is Multiple Criteria Decision Analysis (MCDA). The outcome of this paper is the score aggregation method proposed to be applied in LitIO developed using MCDA approaches.

Keywords: informatics olympiads, programming contests, evaluation, grading, multiple criteria decision analysis.

1. Introduction

The Lithuanian Informatics Olympiad (LitIO) is a state supported algorithmic problem solving competition for students in secondary education. The contestants are given algorithmic tasks and have to solve them in four or five hour contest sessions. They have to design and implement the algorithm in order to solve the task. The task may also require to submit reasoning for algorithm design or a set of test cases. The material submitted for evaluation by a contestant is called *a submission*. After the submission has been evaluated in terms of separate criteria where each criterion is measured according to its own scale (but the same scale for each contestant), the aggregated score has to be calculated so that the submissions can be ranked with respect to other submissions for the same task. Measuring the distance between contestants is also important.

This paper is a continuation of the paper *Improving the Evaluation Model for the Lithuanian Informatics Olympiads* (Skūpienė, 2010), where we elaborated the components of the submission, and the list of evaluation criteria. Each criterion consists of a description, the form of evaluation (manual, automatic) and the scale (boolean, ordinal, ratio).

In this paper we focus on the score aggregation once the submissions have been evaluated in terms of the separate criteria. We approach the problem from the point of view of *Multiple Criteria Decision Analysis* (MCDA). The structure of the paper is the following.

First we will explain the main ingredients of MCDA and show that submission ranking in LitIO falls under the category of problems that can be solved using MCDA techniques. Then we will present the requirements that emerge from the specifics of the problem, and finally we focus on the choice of the MCDA algorithms, suitable for score aggregation in LitIO.

2. Submission Ranking in LitIO as an MCDA Problem

The field of Multiple Criteria Decision Analysis (MCDA) is defined as a collection of formal approaches which seek to take into account multiple criteria in order to help decision makers to explore different decision alternatives (Belton and Stewart, 2003).

Evaluation in LitIO as such can be treated as an MCDA problem, and the work presented in Skūpienė (2010) corresponds to the first stage of the MCDA process, i.e., problem structuring. The outcome of problem structuring is an explicit list of criteria and alternatives. The task that has to be explored in this paper is ranking of submissions once the submissions have been evaluated in terms of separate criteria. Note, that in practice, the overall ranking has to be based on the scores of several tasks. However, in this paper, we limit our research to determining a score for one task only.

Three major roles can be identified in MCDA. They are *decision maker*, *decision analyst* and *stakeholder* (Val, 2002).

The scientific part of LitIO is managed by the scientific committee. The scientific committee is responsible for all the scientific decisions, i.e., approving the syllabus of the contest, designing tasks and tests, approving the evaluation procedure, performing the evaluation, approving ranking and declaring winners. In 2010 the scientific committee of LitIO consisted of 13 members (Scientific Committee of Lithuanian Informatics Olympiads, 2010). The scientific committee is the only decision maker in this context. The role of decision analyst is played by the author of this paper.

The most important stakeholders are interested in programming and algorithmics students in secondary education from all over Lithuania, as well as the community of informatics teachers. This community of stakeholders is affected directly by each decision or change in the evaluation scheme. Scientific committee of LitIO is also a stakeholder, because possible changes in the evaluation scheme might change their working procedures, time spend on task design and evaluation.

Model of relationship between the different roles in the decision analysis process in the problem under consideration is presented in Fig. 1.

There were suggested several ways how to classify MCDA problems (Roy, 1996; Belton and Stewart, 2003). Submission ranking problem is *the ranking problematique* as the final outcome of the evaluation is a ranked list of contestants based on which the awards will be distributed.

Based on another type of classification, the submission ranking problem is *repeated* problem, therefore the focus of this research is on constructing the ranking procedure which could be applied annually in LitIO.

It is a *group decision making* problem, because the role of decision maker is played by the members of LitIO scientific committee and the opinions of all of those members

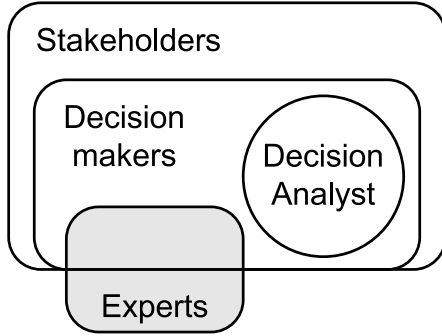


Fig. 1. Model of relationship between different roles in decision analysis in the analysed problem.

who are involved in the evaluation of submissions of a particular task has to be taken into account.

3. Decision Matrix

Once the list of alternatives and criteria for an MCDA problem is determined, the following step is to construct a decision matrix. In this section we present the decision matrix constructed for the submission ranking problem.

Let $A = \{A_1, A_2, \dots, A_n\}$ is a finite set of alternatives (i.e., submissions), $C = \{C_1, C_2, \dots, C_m\}$ is a finite set of criteria (i.e., evaluation criteria), and $P = \{P_1, P_2, \dots, P_q\}$ is a finite set of decision makers (i.e., evaluators).

Assume that all the decision makers are involved into defining relative weights and determining performance of each alternative in terms of each attribute.

Let w_j^k is a relative weight of criteria j ($j = 1, 2, \dots, m$) given by decision maker P_k ($k = 1, 2, \dots, q$).

Let x_{ij}^k is the performance of alternative A_i ($i = 1, 2, \dots, n$) in terms of criteria C_j ($j = 1, 2, \dots, m$) assigned by decision maker P_k ($k = 1, 2, \dots, q$).

Note that if the criterion is measured subjectively, i.e., the decision makers assess the performance of an alternative i in terms of criterion j manually, then the value of x_{ij}^k is linguistic. Otherwise the value of x_{ij}^k is numeric and $x_{ij}^1 = x_{ij}^2 = \dots = x_{ij}^q$.

Then the submission ranking problem can be expressed by the following decision matrix:

$$D^k = (x_{ij}^k)_{n \times m \times q} = \begin{pmatrix} & C_1 & C_2 & \dots & C_m \\ & w_1^k & w_2^k & \dots & w_m^k \\ A_1 & x_{11}^k & x_{12}^k & \dots & x_{1m}^k \\ A_2 & x_{21}^k & x_{22}^k & \dots & x_{2m}^k \\ \dots & \dots & \dots & \dots & \dots \\ A_n & x_{n1}^k & x_{n2}^k & \dots & x_{nm}^k \end{pmatrix}, \quad (1)$$

where $k = 1, 2, \dots, q$.

Note, that the classical MCDA algorithms, assume single decision maker problem, i.e., they assume that $q = 1$.

4. Application of Fuzzy Numbers for Quantifying Linguistic Variables

Some of proposed evaluation in LitIO criteria are measured manually using linguistic variables. Linguistic variables are variables whose values are linguistic terms and not numbers. They are used to express results of subjective qualitative evaluation. Linguistic variables were introduced and described by Zadeh (1975a, b, c). Triangular and trapezoidal fuzzy numbers are used for quantifying linguistic variables.

Next we shortly present the related concepts of fuzzy logic based on Lee (2005), Triantaphyllou (2000).

Fuzzy set. A Fuzzy set is any set that allows its members to have different grades of membership (membership function) in the interval $[0, 1]$, i.e., for any subset \tilde{A} of a universe X it is possible to define a membership function of a fuzzy set: $\mu_{\tilde{A}} : X \rightarrow [0, 1]$.

A crisp set is a separate case of fuzzy set and to make distinctions between crisp and fuzzy sets we will use \tilde{A} notation for fuzzy sets.

Operations on Fuzzy sets (Zadeh, 1965; Lee, 2005; Triantaphyllou, 2000):

- *Negation:* $\mu_{\tilde{A}^c}(x) = 1 - \mu_{\tilde{A}}(x), \forall x \in X,$
- *Union:* $\mu_{\tilde{A} \cup \tilde{B}}(x) = \text{Max} [\mu_{\tilde{A}}(x), \mu_{\tilde{B}}(x)], \forall x \in X,$
- *Intersection:* $\mu_{\tilde{A} \cap \tilde{B}}(x) = \text{Min} [\mu_{\tilde{A}}(x), \mu_{\tilde{B}}(x)], \forall x \in X.$

Fuzzy number. Fuzzy set is called a fuzzy number if the fuzzy set is convex, normalised, its membership function is defined in \mathbb{R} and is piecewise continuous.

Trapezoidal fuzzy number. Trapezoidal fuzzy number is a fuzzy number represented with four points as follows: $\tilde{A} = (a_1, a_2, a_3, a_4)$ and this representation is interpreted in the following way:

$$\mu_{\tilde{A}}(x) = \begin{cases} 0, & x < a_1, \\ \frac{x - a_1}{a_2 - a_1}, & a_1 \leq x \leq a_2, \\ 1, & a_2 \leq x \leq a_3, \\ \frac{a_4 - x}{a_4 - a_3}, & a_3 \leq x \leq a_4, \\ 0, & x > a_4. \end{cases} \quad (2)$$

When $a_2 = a_3$, the trapezoidal number coincides with a *triangular fuzzy number*.

There were created many conversion scales for transforming linguistic terms into fuzzy numbers. Chen *et al.* (1992) proposed eight conversion scales with different numbers of linguistic terms which are commonly used. An example of pretty standard in fuzzy set theory nine-item scale is presented in Table 1 and Fig. 2 (Sule, 2001). The choice of a concrete scale from the available ones is intuitive and left to the responsibility of the decision maker. Note that the same linguistic term in different conversion scales can have different crisp values.

Thus, all the linguistic terms, are converted to fuzzy numbers using the chosen scales, and after this the decision matrix will contain only numeric (crisp or fuzzy) values. In this

Table 1
Weights of a trapezoidal distribution of a linguistic scale

Item of linguistic scale	Numerical weights
Very poor (VP)	(0, 0, 0, 0.2)
Between poor and very poor (BPV)	(0, 0.2, 0.2, 0.4)
Poor (P)	(0, 0.2, 0.2, 0.4)
Between poor and fair (BPF)	(0, 0.2, 0.5, 0.7)
Fair (F)	(0.3, 0.5, 0.5, 0.7)
Between fair and good (BFG)	(0.3, 0.5, 0.8, 1)
Good (G)	(0.6, 0.8, 0.8, 1)
Between good and very good (BGV)	(0.6, 0.8, 0.8, 1)
Very good (VG)	(0.8, 1, 1, 1)

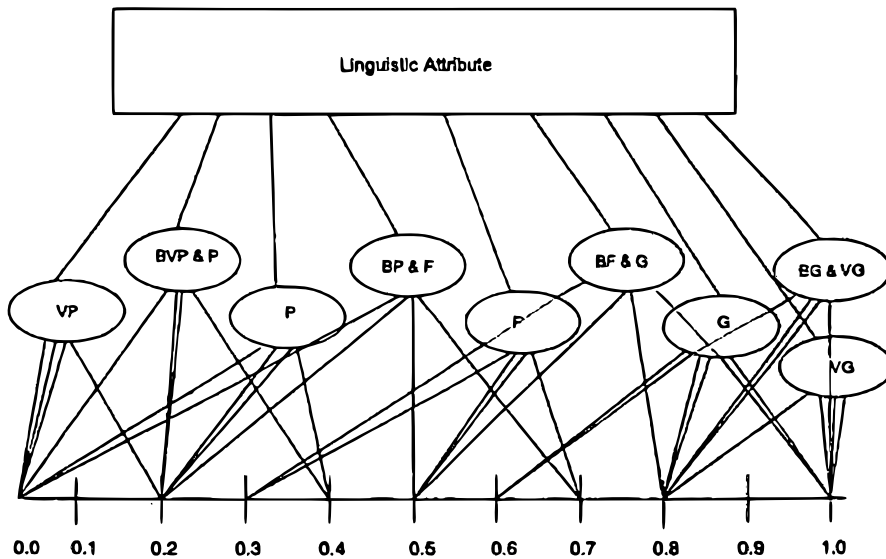


Fig. 2. Trapezoidal fuzzy numbers are used to quantify nine-item linguistic scale (Sule, 2001).

paper we will not suggest the concrete scales, because the scales are chosen intuitively and we believe that the jury members also have to be involved in the decision. Only after the piloting of the evaluation scheme it might be possible to make a final decision about the scales.

5. Submission Ranking Problem Constraints

The decision context of our problem is rather specific. The problem belongs to the ranking problematique category and is a group decision making problem. Moreover, the chosen

method will have to be applied in an educational informatics contest situation. Therefore it is highly important that the approach would be *accepted by the community* of informatics contests. Belton and Stewart (2003) emphasize that the ability to explain the chosen approach to a variety of backgrounds is an important factor in the choice of MCDA approach.

The score aggregation contains parts which are revealed to the contestants, but it also contains the hidden parts. For example, the scores assigned by individual jury members during manual evaluation are not revealed to the contestants, only the aggregated score is. We emphasise that the parts of the scoring function which are revealed to the contestants must be easily understandable and transparent. While more complicated techniques could be applied to the hidden parts.

It must be noted that our problem is a repeated problem. This means that the process of ranking submissions will have to be repeated each time a LitIO contest session takes place. Though on different submissions possibly of different contestants. Therefore it is very important to achieve that the stakeholders would accept and understand the method.

Even though the problem is described as ranking problematique, it is not enough to present ranking to the contestants. The jury (during medal allocation procedure) and the contestants are interested not only in the position in the ranking table, but also in the score differences among a group of contestants.

It is commonly accepted in LitIO that a score aggregation function mapping the performances for separate attributes (groups of criteria) into real numbers is defined and announced to the contestants in advance.

Therefore we will focus on MCDA approaches which foresee defining score aggregation function and partial score functions, inducing a ratio scale, and the ranking is made after the aggregated scores for each alternative have been calculated.

6. Choice of MCDA Approach

Many different MCDA approaches are presented and categorised (Triantaphyllou, 2000; Kahraman, 2008; Chen *et al.*, 1992; Carlsson and Fullér, 1996; Belton and Stewart, 2003). Instead of focusing of separate MCDA methods, we will first look at the major families of MCDA methods. Belton (2003) distinguish three major families of MCDA approaches:

- *Value measurement theory* (Keeney and Raiffa, 1976). The main idea of this approach is to construct a value function which would associate each alternative with a real number in order to produce ranking of alternatives. The main idea of this theory correspond the intentions and reasoning for our problem. Therefore we will include it for further consideration.
- *Satisficing (or Goal programming)* (Simon, 1976). This approach instead of creating one value function operates on partial value functions. By partial value function we understand value function which maps performance of alternatives in terms of a certain criteria to real number. The main idea of the approach is that the most important criterion is identified and the acceptable level of it is determined. Then

the alternatives are eliminated until all the remaining alternatives achieve the acceptable level. At this point the second most important alternative together with its satisfactory level is identified. The alternatives which do not reach satisfactory level of the second criteria are eliminated again.

This approach is not suitable for our problem as it does not assume score aggregation at all.

- *Outranking* (Roy, 1996). Outranking methods also operate with partial value functions and involve pair-wise comparisons of alternatives. An alternative is dominated by another alternative if the another alternative performs better in terms of one or more criteria and equals in the remaining criteria. The concept of *outranking* is introduced. The outranking relationship of two alternatives describes that even though the two alternatives do not dominate each other mathematically, the decision maker accepts the risk of regarding one alternative almost surely better than the other.

We consider this approach also unacceptable in our situation because it again deals with preferences in terms of separate criteria and does not foresee score aggregation using single value function. The concept of outranking, i.e., allowing the decision maker to take the risk of considering one alternative better than the other is not acceptable in a contest community where scoring is extremely sensitive issue.

Out of three major MCDA families, only one foresees construction value function, which is required in ranking submissions problem as well. Therefore further we focus on algorithms of value measurement theory.

Besides main families of MCDA approaches, fuzzy logic is often considered to be applied for MCDA problems. Fuzzy logic is used in group decision making which is our case. However, fuzzy logic is not a separate methodology, but a tool that can be applied within other MCDA approaches including the ones described above. Therefore we assume that fuzzy logic might be applicable in case of this problem and we will look at the concepts of fuzzy logic as well.

7. Choice of Value Measurement Theory Method

Value measurement theory was mainly started by Keeney and Raiffa (1976). More on it can be found in Roberts (1979), French (1988).

The main idea of this theory is that a real number (“value”) is associated with each alternative in order to produce ranking of alternatives. *The value function* is defined as a function assigning a non-negative number to each alternative indicating the desirability (or preference) of the alternative.

The value function has to satisfy the following requirements: an alternative A_{i1} is preferred to alternative A_{i2} ($A_{i1} \succ A_{i2}$) if and only if $V(A_{i1}) > V(A_{i2})$; the alternatives are indifferent if and only if $V(A_{i1}) = V(A_{i2})$. Note, that the value function must induce complete order.

Value measurement approach introduces partial value functions $v_j(A_i)$. They are constructed for separate criteria and partial value functions holds the essential features (i.e., induces complete order) of value function in terms of separate criteria.

There were developed several value measurement theory algorithms and the most popular ones are *Weighted Sum Model* and *Weighted Product Model*. We would also assign *Topsis* algorithm (we will present it later) to the same category of algorithms. Note that those algorithms are constructed for single decision maker problems.

Weighted Sum Model (WSM) is most commonly used method for single decision maker problems (Triantaphyllou, 2000). It can be described using the following formula:

$$V(A_i) = \sum_{j=1}^m w_j v_j(A_i), \quad (3)$$

where $i = 1, 2, \dots, n$ and $j = 1, 2, \dots, m$.

One of the reasons of wide acceptance of this model is its simplicity, i.e., it can be easily explained by the decision makers to a variety of backgrounds (Belton and Stewart, 2003).

Note that the requirement *preferential independence* has to be satisfied so that the WSM model could be applied (Belton and Stewart, 2003). Suppose that two alternatives A_{i1} and A_{i2} differ only on a set of criteria $R \subset C$ (R is a proper subset of C) and values of partial functions are equal on all other criteria. Then it is possible to decide the relationship of A_{i1} and A_{i2} (i.e., $A_{i2} \succ A_{i1}$ or $A_{i1} \succ A_{i2}$ or $A_{i1} \sim A_{i2}$) knowing their performances on criteria from R only, i.e., irrespective of values of their performances on all the other criteria.

However among the criteria there are several dependent criteria, e.g., quality of programming style is related either to the performance of an algorithm-code complex or to its efforts to solve the task (Skūpienė, 2010). Thus partial independence of criteria is violated.

We suggest that this does not eliminate WSM from applying it for score aggregation in LitIO. WSM still can be applied for aggregating those criteria that are preferentially independent. Special functions have to be introduced for aggregating scores of dependent criteria.

Another requirement is to use the same scale of measurement for all the criteria. Performance of submissions are measured using different scales. However we intend to unify the scales by constructing the corresponding partial value functions.

WSM can be potentially applied for score aggregation in LitIO, though the above mentioned condition has to be observed.

Weighted Product Model (WPM). WPM can be described using the following formula:

$$V(A_i) = \prod_{j=1}^m [v_j(A_i)]^{w_j}, \quad (4)$$

where $i = 1, 2, \dots, n$ and $j = 1, 2, \dots, m$.

There have been suggested arguments that preferences are often perceived in ratio scale terms therefore product is more natural than sum (Lootsma, 1997; Triantaphyllou, 2000). The consequence of trade an additive approach into multiplicative approach is that partial value functions have to satisfy ratio scale properties instead of interval scale properties (Triantaphyllou, 2000). The WSM presented above uses partial value functions which can solve this issue.

Simplicity of the approach is a high priority in the choice of score aggregation algorithm. We conclude that WSM algorithm would be more suitable than WPM as it is simpler and better understandable to the wide audience and otherwise they seem to be identical in terms of the problem under consideration.

Topsis (Technique for Order Preference by Similarity to Ideal Solution) (Triantaphyllou, 2000; Saghafian and Hejazi, 2005). We did not find it explicitly stating that Topsis belongs to Value measurement theory approaches, nor to other specific family of MCDA approaches. However, as it involves calculating value of closeness coefficient and ranking based on the values of the coefficient, we suggest that it is appropriate to consider it here.

Topsis introduces concepts of hypothetical solutions, i.e., positive ideal solution and negative ideal solution. The positive ideal solution is calculated as a function from the best performance values of the concrete decision matrix in terms of each criteria: $A^* = \{v_{1*}, v_{2*}, \dots, v_{m*}\}$, where $v_{j*} = \max_i(v_j(A_i))$, $i = 1, \dots, n$.

The negative ideal solution is calculated as a function from the worst performance values in terms of each criteria: $A^- = \{v_{1-}, v_{2-}, \dots, v_{m-}\}$, where $v_{j-} = \min_i(v_j(A_i))$, $i = 1, \dots, n$.

For each alternative the Euclidean distance from ideal positive solution and ideal negative solution is calculated: $S_{i*} = \sqrt{\sum_{j=1}^m (v_j(A_i) - v_{j*})^2}$, and $S_{i-} = \sqrt{\sum_{j=1}^m (v_j(A_i) - v_{j-})^2}$. Finally, the relative closeness coefficient to the ideal positive solution is calculated: $F_{i*} = \frac{S_{i-}}{S_{i*} + S_{i-}}$.

The alternatives are ranked based on the value of relative closeness coefficient to the ideal solution of each alternative. This method from mathematical point of view is interesting and appealing however, it gives in to WSM due to the simplicity of the latter.

Moreover, score of one submission is dependant upon the quality of the submissions. There were cases where such approach was applied in large informatics contests. However, in LitIO contestants also compete in small groups and there are cases where just few (i.e., less than 10) submissions per task are presented. If the score is dependant upon the submissions in such case, then it might become too biased.

After looking at several value measurement theory associated methods, we came to that conclusion, that as simplicity and the ability of wide audience to accept the evaluation scheme plays significant role in the choice of approaches, WSM approach suits best to solving evaluation in LitIO problem. Though certain requirements have to be observed. We did not find any evidence that other methods would be more suitable than WSM.

We decided on the score aggregation method, however the classical WSM assumes single decision maker and therefore we have to look for the extension of WSM to be applicable for group decision problems.

8. Group Decision Making

Group Decision Making (GDM) can be defined as decision making process based on the opinions of several individuals. The goal of GDM is to arrive to the satisfactory for the group solution, rather than to the best solution which almost does not exist (Lu *et al.*, 2007). There are available various methods for group decision making from mathematical to psychological and social.

Among MCDA approaches explicitly meant for solving group decision making problems there are techniques which foresee negotiation theory, working with group dynamics, etc. References to that can be found (Carlsson and Fullér, 1996; Lu *et al.*, 2007). Those approaches have been experienced in LitIO many times. Investigation of their suitability in LitIO evaluation problem would require much investigations from other sciences, in particular management and psychology. For example, most of meetings are conducted on-line (as members of the scientific committee are associated with different universities in different cities and even countries), some members are reluctant to discuss issues on-line, less experienced tend to vote as more experienced members, etc. These aspects should have been investigated if the above mentioned direction was taken.

Our choice is to focus on mathematical group decision making methods which assume eliciting concrete information from decision makers and using it in a mathematical algorithm, but do not require interaction and negotiation between decision makers.

There are different ways how to implement group decision making. Much references can be found (Rao, 2007; Lu *et al.*, 2007). Many common GDM methods (e.g., *authority rule*, *majority rule*, *negative minority rule*) are not suitable because they are intended for *the choice problematique* (i.e., determining the best alternative), but not for *the ranking problematique* problems.

Lu *et al.* (2007) distinguishes three factors which influence GDM:

The weights of the decision makers. Among the decision makers there might be those who play more important roles in the decision making. In this case, the decision makers should be assigned different weights and that should be reflected in the group decision making process.

Weights of criteria. The decision makers may have different views, attitudes, experience and therefore propose different weights to the criteria.

Preferences of decision makers for alternatives. If the performance of an alternative is evaluated subjectively, then different decision makers can have different understanding, different experiences and evaluate performance of the same alternative in a different way.

It is common in GDM that the weight of a decision maker, the suggested weights for evaluation criteria, and the performances of alternatives suggested by the decision makers are expressed by linguistic terms, since linguistic terms reflect uncertainty, inaccuracy, and fuzziness of the decision makers (Lu *et al.*, 2007). We also assume that the information, provided by each decision maker is consistent and non-conflicting.

9. Score Aggregation Method for Submission Ranking

We already concluded that WSM approach suits best for submission ranking problem. We were looking for an extension of WSM to GDM, such that it would allow fuzzy numbers in the decision matrix, however would use crisp number for partial scores for the attributes and for the final ranking, i.e., its public parts would be acceptable by the community of LitIO.

Many fuzzy GDM algorithms (e.g., *an intelligent FMCGDM method* (Lu *et al.*, 2007) or the one described in Sule (2001)) assume aggregating fuzzy numbers and only then deriving the final ranking. There was performed a systematic and critical study of the existing fuzzy MCDA methods. It was arrived to the conclusion, that the majority of currently existing fuzzy MCDA approaches involve complicated calculations, require all the elements of decision matrix to be presented in a fuzzy format (though some of them might be crisp), and are not suitable for solving problems with more than ten alternatives associated with more than ten criteria (Rao, 2007; Chen *et al.*, 1992).

The method presented by Chen *et al.* (1992) is considered to be the one which avoids the above mentioned problems (Rao, 2007; Zhang, 2004). It consists of the following phases:

- linguistic terms (if such are used) are converted to fuzzy numbers;
- fuzzy numbers are converted into crisp scores;
- classical MCDA approaches, which assume crisp values, are applied.

Now we have to find a classical GDM method which assumes crisp matrix. Such method is *The group decision support* algorithm suggested by Csáki (1995).

Therefore we have to combining the group decision support algorithm with the approach of Chen *et al.* (1992). Thus we obtain a GDM algorithm suitable to apply in LitIO evaluation problem. Further we use the notations introduced in the Section 3.

The linguistic terms are converted to fuzzy numbers as it was described in Section 4. The crisp score of a fuzzy number \tilde{A} is calculated in the following way. First there are defined two functions $\mu_{\max}(x)$ and $\mu_{\min}(x)$:

$$\mu_{\max}(x) = \begin{cases} x, & 0 \leq x \leq 1, \\ 0, & \text{otherwise,} \end{cases} \quad (5)$$

$$\mu_{\min}(x) = \begin{cases} 1 - x, & 0 \leq x \leq 1, \\ 0, & \text{otherwise.} \end{cases} \quad (6)$$

Then the left and the right scores of \tilde{A} and the two functions are defined as:

$$\mu_L(\tilde{A}) = \text{Sup}_x[\mu_{\tilde{A}}(x) \cap \mu_{\min}(x)], \quad (7)$$

$$\mu_R(\tilde{A}) = \text{Sup}_x[\mu_{\tilde{A}}(x) \cap \mu_{\max}(x)]. \quad (8)$$

Here *Sup* stands for the least upper bound. The total score crisp score of the fuzzy number \tilde{A} is defined as:

$$\mu_T(\tilde{A}) = (\mu_R(\tilde{A}) + 1 - \mu_L(\tilde{A}))/2.$$

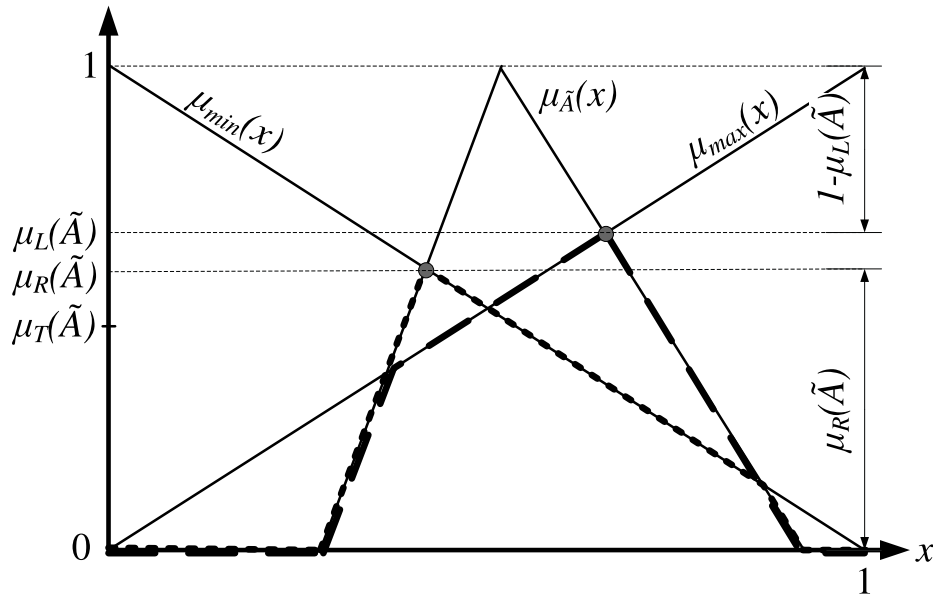


Fig. 3. Conversion of triangular fuzzy number to crisp value.

Table 2
Calculating crisp values of the nine-item linguistic scale given in Table 1

Item of linguistic scale	Fuzzy number \tilde{A}	$\mu_R(\tilde{A})$	$\mu_L(\tilde{A})$	$\mu_T(\tilde{A})$
Very poor	(0, 0, 0, 0.2)	1	0.17	0.08
Between poor and very poor	(0, 0.2, 0.2, 0.4)	0.83	0.33	0.25
Poor	(0, 0.2, 0.2, 0.4)	0.83	0.33	0.25
Between poor and fair	(0, 0.2, 0.5, 0.7)	0.83	0.58	0.38
Fair	(0.3, 0.5, 0.5, 0.7)	0.58	0.58	0.50
Between fair and good	(0.3, 0.5, 0.8, 1)	0.58	0.83	0.63
Good	(0.6, 0.8, 0.8, 1)	0.33	0.83	0.75
Between good and very good	(0.6, 0.8, 0.8, 1)	0.33	0.83	0.75
Very good	(0.8, 1, 1, 1)	0.17	1.00	0.92

Conversion of a fuzzy number to crisp value is illustrated in Fig. 3.

The values of the nine-item linguistic scale presented in Table 1 converted to crisp values are presented in Table 2. Note that the same linguistic term in different conversion scales can have different crisp values.

The algorithm for converting fuzzy numbers to crisp values might be hardly understandable to the wide audience, however its application will remain invisible for the contestants. It will only be applied for dealing with group decisions and linguistic evaluation. If a criterion requires manual evaluation, the linguistic scores and the scores of individual jury members are never revealed to the contestants, just the aggregated score for the crite-

tion is announced. Thus, if fuzzy techniques are used to aggregate scores of several jury members, they remain behind the curtains and do not become the source of discussions and doubts for the contestants.

The final step is to apply the *group decision support* algorithm (Csáki *et al.*, 1995) to the crisp decision matrix.

First the aggregated group weights for each criterion are calculated:

$$w_j = \frac{\sum_{k=1}^t w_j^k p_k}{\sum_{k=1}^t p_k}, \quad j = 1, 2, \dots, m. \quad (9)$$

The values of partial value functions of performance of each alternative in terms of each criterion are calculated in a similar way:

$$v_j(A_i) = \frac{\sum_{k=1}^t v_j^k(A_i) p_k}{\sum_{k=1}^t p_k}. \quad (10)$$

The total aggregated values for each alternative are calculated in the following way:

$$v(A_i) = \frac{\sum_{j=1}^m v_j(A_i) w_j}{\sum_{j=1}^m w_j}. \quad (11)$$

Based on the calculated values, the ranking of the alternatives is performed.

Conclusions

In this paper we proposed to combine the group decision support algorithm combined with score aggregation method to be applied during evaluation process in Lithuanian Informatics Olympiad. The method takes into account linguistic values (outcome of manual evaluation) and multiple decision makers (members of the scientific committee).

Even though MCDA theory is acceptable from the scientific point of view, there arise many difficulties with its application in practice because the stakeholders feel reluctant and sensitive about the application of complicated formula to sensitive issues (in this case score aggregation).

The most important requirements to the score aggregation method were the understandability and acceptability to the wide audience of the public (i.e., disclosed to the contestants) parts of it. Another important requirement was use of value function. As a result of these requirements, we spent time on looking for a suitable method that would fulfill all the problem specific requirements, rather than analysing several equally possible options. The paper reveals how we arrived to the suggested score aggregation method.

The next step would be piloting the suggested score aggregation scheme and investigating how it will be accepted in LitIO.

References

- Belton, V., Stewart, T.J. (2003). *Multiple Criteria Decision Analysis: An Integrated Approach*. Kluwer Academic Publishing, Boston.
- Carlsson, C., Fullér, R. (1996). Fuzzy multiple criteria decision making: recent developments. *Fuzzy Sets and Systems*, 78(2), 139–153.
- Chen, S.J., Hwang, C.L., Hwang, F.P. (1992). Fuzzy multiple attribute decision making: methods and applications. In *Lecture Notes in Economics and Mathematical Systems*, Vol. 375. Springer-Verlag, Berlin, Germany.
- Csáki, P., Rapcsák, T., Turchányi, P., Vermes, M. (1995). Research and development for group decision aid in Hungary by WINGDSS, a Microsoft Windows based group decision support system. *Decision Support Systems*, 14, 205–221.
- French, S. (1988). *Decision Theory: An Introduction to the Mathematics of Rationality*. Ellis Horwood, Chichester.
- Kahraman, C. (2008). *Fuzzy multi criteria decision making. Theory and applications with recent developments. Optimization and its Applications*, 16. Springer.
- Keeney, R.L., Raiffa, H. (1976). *Decisions with Multiple Objectives: Preferences and Value Tradeoffs*. John Wiley & Sons.
- Lee, K.H. (2005). *First Course on Fuzzy Theory and Applications*. Springer.
- Lootsma, F.A. (1997). *Fuzzy Logic for Planning and Decision Making*. Kluwer.
- Lu, J., Zhang, G., Ruan, D. (2007). *Multi-Objective Group Decision Making: Methods, Software and Application with Fuzzy Set Techniques*. Series in Electrical and Computer Engineering.
- Rao, R.V. (2007). *Decision Making in the Manufacturing Environment*. Springer.
- Roberts, F.S. (1979). *Measurement Theory with Applications to Decision Making, Utility and the Social Sciences*. Addison-Wesley, London.
- Roy, B. (1996). *Multicriteria Methodology for Decision Aiding*. Kluwer Academic Publishers, Dordrecht.
- Saghafian, S., Hejazi, S.R. (2005). Multi-criteria group decision making using a modified fuzzy *topsis* procedure. In: *Computational Intelligence for Modeling, Control and Automation*. IEEE Proceedings.
- Scientific Committee of Lithuanian Informatics Olympiads* (2010).
- Simon, H.A. (1976). *Administrative Behavior*. The Free Press, New York.
- Skūpienė, J. (2010). Improving the evaluation model for the lithuanian informatics olympiads. *Informatics in Education*, 9(1), 141–158.
- Sule, D.R. (2001). *Logistics of Facility Location and Allocation*. Marcel Dekker, New York, Basel.
- Triantaphyllou, E. (2000). *Multi-Criteria Decision Making Methods: A Comparative Study*. Kluwer Academic Publishers.
- Value Tree Analysis* (2002). Multiple criteria decision analysis E-learning site created in the EU project OR-World by System Analysis Laboratory of Helsinki University of Technology.
http://www.mcda.hut.fi/value_tree/theory/.
- Zadeh, L.A. (1965). Fuzzy sets. *Information and Control*.
- Zadeh, L.A. (1975a). The concept of a linguistic variable and its application to approximate reasoning. I. *Information Sciences*, 8(3), 199–249.
- Zadeh, L.A. (1975b). The concept of a linguistic variable and its application to approximate reasoning, II. *Information Sciences*, 8(43), 301–357.
- Zadeh, L.A. (1975c). The concept of a linguistic variable and its application to approximate reasoning, III. *Information Sciences*, 9(1), 43–80.
- Zhang, W. (2004). Handover decision using fuzzy MADM in heterogeneous networks. In: *Wireless Communications and Networking Conference, WCNC, IEEE Vol. 2*. 653–658.

J. Skūpienė is a younger research fellow in the Informatics Methodology Department in the Institute of Mathematics and Informatics Vilnius University. She is a member of the Scientific Committee of National Olympiads in Informatics since 1994 and a Lithuanian team leader in International Olympiads in Informatics since 1996. For a few years she was director of studies of Young Programmers' School by Correspondence. Since 2004 she has been a coordinator of Informatics section in the National Academy of Students. She is author/co-author of three books on algorithmic problems for informatics contests. Her research interests include informatics and algorithmic contests, teaching informatics and computer science in secondary education.

Informatikos olimpiados rezultatų skaičiavimas naudojant daugiakriterinį sprendimų metodą

Jūratė SKŪPIENĖ

Lietuvos mokinių informatikos olimpiados yra algoritmų kūrimo ir programavimo konkursas, skirtas mokiniams. Kiekvienas pateiktas sprendimas yra vertinamas pagal keletą kriterijų, o kiekvienas kriterijus matuojamas pagal atskirą skalę. Vertinimas pagal vienus kriterijus atliekamas automatizuotai, pagal kitus – ne automatinio būdu. Vertina keli vertintojai, kurie parenka vertinimo kriterijų svorius bei atlieka ne automatinį vertinimą. Tačiau kiekvienas vertintojas gali turėti skirtingą požiūrį į tai, koks turėtų būti konkretaus kriterijaus svoris ir koks turėtų būti konkretaus sprendimo įvertis vieno ar kito kriterijaus atžvilgiu. Straipsnyje, taikant daugiakriterių sprendimų teoriją, sprendžiamas klausimas kaip gauti galutinius agreguotus vertinimui pateiktų darbų įverčius. Taip pat analizuojama vertės matavimo teorija ir nagrinėjamas problemai spręsti siūlomas svorinio sumavimo metodo išplėtimas, kuris gautas sujungus grupinių sprendimų palaikymo algoritmą su S.J. Chen pasiūlytu metodu.