

Multiple Criteria Decision Methods in Informatics Olympiads

Jūratė SKŪPIENĖ

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

Abstract. Multiple Criteria Decision Making (MCDM) is applied in a variety of areas, including education. Informatics Olympiads, problem solving contests for high school students, is the area where MCDM methods can also be applied. The case of the Lithuanian Informatics Olympiad is analysed in this paper. There are several aspects occurring while maintaining the contest that requires decision making. 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. Thus we get a problem: how to calculate the aggregated score for whole submission in the above mentioned situation. Another similar problem is making decision on national team selection for other international contests where each candidate is evaluated in terms of several criteria. 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 field of multiple criteria decision analysis (MCDA) is also termed as a multiple criteria decision aid or multiple criteria decision making (MCDM). Its target is to help reach a consensus and compromises between conflicting goals (i.e., multiple criteria) in complex problems.

In real life it is unusual that the problem is presented to the analyst in a form of a clearly defined set of alternatives and criteria (Belton and Stewart, 2003). Problems might be complex and confusing and they typically involve a wide range of criteria that need to be considered. They might involve conflicting criteria, the conflicts between different stakeholders about the importance of criteria in making a decision. It might even be required to define criteria as they are not clear at the initial stage of the problem. The general goal of MCDA is to assist individual or groups of decision makers to choose the best alternative.

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).

Potential problems that MCDA can be applied to solve come from a variety of areas like business, medicine, banking, marine industry, bioinformatics, public policies or education (Aruldoss *et al.*, 2013).

Education is one of the areas where MCDA can be widely applied. These are learning content and learning software evaluation problems (Kurilovas and Serikoviene, 2010), higher education decision making problems, (resource allocation, performance management, budgeting and scheduling) (Ho *et al.*, 2006), using MCDA for accreditation in order to evaluate IT skills and qualifications (Siskosa *et al.*, 2007), evaluating factors that determine the quality of higher education (Tsinidou *et al.*, 2010), evaluation of the quality of e-learning systems (Tzenga *et al.*, 2007) and educational websites (Shee and Wang, 2008), pedagogical evaluation of teachers (Filipe *et al.*, 2015), evaluating quality of learning objects (Kurilovas *et al.*, 2011).

The majority of research of application of MCDA in education that we discovered was related to the evaluation of quality of various educational factors or tools. However, we noticed that the choice of MCDA approaches highly depends upon the category of the problem under consideration.

Four broad categories of MCDA problems have been proposed (Roy, 1996):

- **The choice problematique.** Problems fall into this category if there is a need to make a choice from a set of alternatives. However the set of alternatives might be either finite or infinite.
- **The sorting problematique.** In this case the given alternatives have to be sorted into several categories, such as “definitely acceptable”, “possibly acceptable”, “definitely unacceptable”.
- **The ranking problematique.** The alternatives have to be ranked in some order of preference.
- **The description problematique.** Possible alternatives and their consequences have to be described formally in a systematic way so that the decision makers could evaluate the alternatives.

Variations or amendments to this classification are also possible (Belton and Stewart, 2003).

Another classification of MCDA problems is *one-off* versus *repeated* problems. In some cases, a decision has to be made only once as the problem is unique. This is a one-off problem and the process is oriented towards arriving at a specific decision. In the case of repeated problems the same problem is recurring a few times or periodically. Then MCDA is oriented towards creating a procedure to be used in decision making.

An MCDA problem can also be classified either as a *single decision making* or *group decision making* problem. In the case of a group decision making problem, several decision makers are involved and they can have different values and opinions how to address the problem. In order to approve the decision, the consensus and compromise among different decision makers has to be reached.

Different authors suggest different stages of the MCDA process. (Val, 2002) proposes a scheme consisting of four stages in particular, problem structuring (decomposed into five sub-stages), preference elicitation, recommended decision, and sensitivity analysis. (Oberti, 2004) suggests four stages of the MCDA process, i.e., beginning of the study, evaluation of actions, multiple criteria modelling, multiple criteria processing, and recommendations.

Each stage consists of two or three sub-stages. (Belton and Stewart, 2003) offer three stages: problem identification and structuring, model building, and using a model to inform and challenge thinking. The scheme based on (Belton and Stewart, 2003) is presented in Fig. 1.

These stages reflect a variety of approaches to MCDA, however, they confirm that an extensive problem analysis and structuring are vital before mathematical algorithms can be applied. In all those approaches the stages are iterative and interactive, i.e., they foresee a return to previous stage, review and update its outcome.

Even though mathematical MCDA algorithms help to arrive at some acceptable alternative, many authors emphasize that MCDA cannot be used to arrive at the “right” answer and it cannot provide a fully objective analysis and totally eliminate subjectivity (Belton and Stewart, 2003). The process of MCDA is emphasised more than the decision it helps to arrive at (Keeney and Raiffa, 1976; Roy, 1996; Zeleny, 1982). The process involves not only the application of mathematical algorithms to come up to the final decision, but also learning about the problem, identifying the key concerns, priorities, uncertainties, values, exploring and generating different alternatives. This should lead to better explainable and justifiable decisions.

2. Submission (Contestant) Ranking in LitIO as an MCDA Problem

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

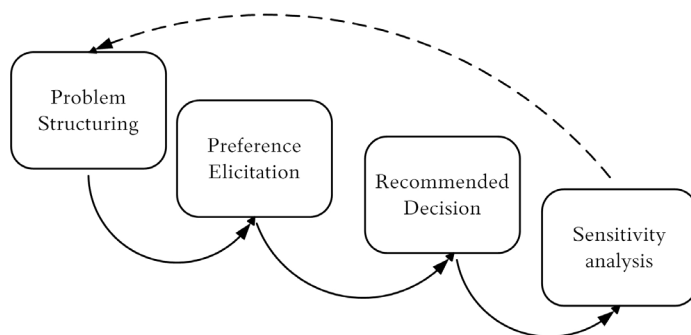


Fig. 1. Basic stages of the MCDA process.

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.

If the contest also serves as team selection event, then the decision has to be made which contest participants will be invited to represent country in a regional or international contest. In the long term practice several criteria for this selection were used and the decision is made by the Scientific Committee of LitIO.

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.

Team selection in LitIO can also be treated as an MCDA problem. The list of selection criteria, as well as the procedure how to evaluate a contestant against each criterion has been decided many years ago. The criteria are: score of the final round of LitIO in the current year; score of the regional contest in the current year (if applicable); awards received in regional and international competitions (there's been approved a concrete list of such competitions); competitor's grade in the current year. There is a consensus over that among the decision makers.

However the criteria are not directly comparable and the understanding of the goals of the team selection varies. Therefore the discussions and the search for the algorithm for the team selection continues. The overall problem - determining the ranking based on the criteria can also be treated as an MCDA problem.

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 and selecting teams to represent Lithuania in regional and international contests. In 2015 the scientific committee of LitIO consisted of 22 members (Sci, 2015). 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 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. The scientific committee of LitIO is also a stakeholder, because possible changes in the evaluation scheme might change their working procedures, time spent on task design and evaluation.

The model of relationship between the different roles in the decision analysis process in the problem under consideration is presented in Fig. 2.

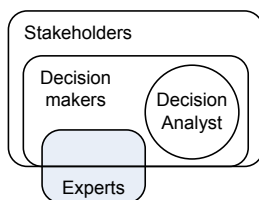


Fig. 2. Model of relationship between different roles in decision analysis in the analysed problem.

There were suggested several ways how to classify MCDA problems (Belton and Stewart, 2003; Roy, 1996). Submission (contestant) 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 or team selected.

Based on another type of classification, the submission ranking problem is *a 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 the LitIO scientific committee and the opinions of all of those members 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 both previously described situation, i.e. for the submission ranking problem and for the contestant ranking problem.

Let $A = \{A_1, A_2, \dots, A_n\}$ be a finite set of alternatives (i.e., submissions or contestants), $C = \{C_1, C_2, \dots, C_m\}$ be a finite set of criteria (i.e., evaluation criteria or team selection criteria), and $P = \{P_1, P_2, \dots, P_q\}$ be a finite set of decision makers (i.e., scientific committee members).

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

Let w_j^k be 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 be 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 $x_{ij}^1 = x_{ij}^2 = \dots = x_{ij}^q$.

In the case of team selection problem the value of x_{ij}^k is always numeric.

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

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

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

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

4. Application of Fuzzy Numbers for Quantifying Linguistic Variables

Some of the 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 (Lee, 2005; Triantaphyllou, 2000; Zadeh, 1965):

- 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. A 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. A 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*.

Many conversion scales have been created 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 pretty standard fuzzy set theory nine-item scale is presented in Table 1 and Fig. 3 (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 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.

Table 1
Weights of a trapezoidal distribution of a linguistic scale (Sule, 2001)

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)

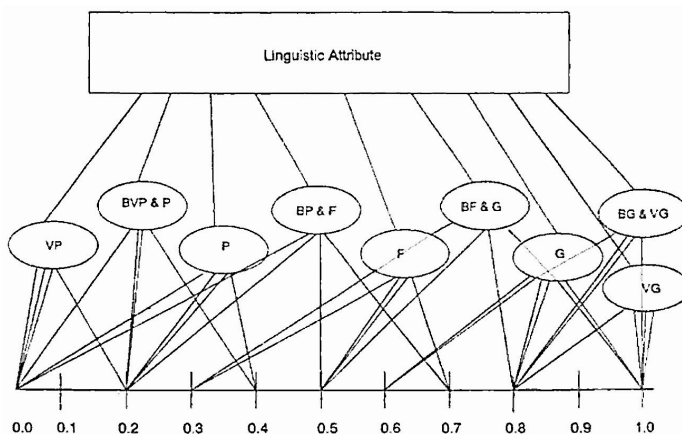


Fig. 3. Trapezoidal fuzzy numbers are used to quantify nine-item linguistic scale (Sule, 2001).

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 to approach to a variety of backgrounds is an important factor in the choice of MCDA approach.

The score aggregation and team selection procedures 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. 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 and contestants 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 in (Belton and Stewart, 2003; Carlsson and Fullér, 1996; Chen *et al.*, 1992; Kahraman, 2008; Triantaphyllou, 2000). Instead of focusing of separate MCDA methods, we will first look at the major families of MCDA methods. (Belton and Stewart, 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 mean a 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 other 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 a sensitive issue.

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

Besides the 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 (Keeney and Raiffa, 1976). More on it can be found in (French, 1988; Roberts, 1979).

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 ($A_{i1} \sim A_{i2}$) 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 hold the essential features (i.e. induces complete order) of value functions 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 the 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 \mathbb{R} only, i.e. irrespective of values of their performances on all the other criteria.

However among the submission evaluation 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 the 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.

In the team selection problem the partial independence requirement is fully satisfied.

Another requirement is to use the same scale of measurement for all the criteria. Performance of submissions as well as performance of contestants 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.

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

For the team selection process WPM can be considered as an option, because the overall number of stakeholder is much smaller. Usually there are not more than 10 candidates among which the selection is performed.

Topsis (*Technique for Order Preference by Similarity to Ideal Solution*) (Saghafian and Hejazi, 2005; Triantaphyllou, 2000). We did not find it explicitly stated 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:

$$A^* = \{v_{1*}, v_{2*}, \dots, v_{m*}\}, \text{ where } v_{j*} = \max_i (v_j(A_i)), i = 1, 2, \dots, n.$$

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 $v_{j-} = \min_i (v_j(A_i))$, $i = 1, 2, \dots, n$.

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 the 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, the 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, LitIO contestants also compete in small groups and there are cases where just few (i.e. fewer than 10) submissions per task are presented. If the score is dependant upon the submissions in such case, then it might become too biased.

In case of team selection, the number of alternatives is allways low (i.e. 5 to 12), however we might consider this method, because the contestant evaluation scores play different role in team selection than in submission evaluation. In submission evaluation the contestants are interested in the absolute score, i.e. how good the contestant is performing in terms of a concrete task. However in the team selection process the main question that matters is how good the contestant is in comparison to other contestants.

After looking at several value measurement theory associated methods, we came to the conclusion, that as simplicity and the ability of wide audience to accept the evaluation scheme plays significant role in the choice of approaches, the 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.

The situation is different with team selection problem. The overall number of stakeholders is much lower (not more than 12 contestants, their teachers and parents), relative ranking is more important than performance in terms of a separate criterion, hidden and revealed parts of the selection process differ form that of submission evaluation. Therefore at this stage all the three methods, i.e. WSM, WPM and Topsis, can be presented for further consideration.

We decided on the score aggregation methods, however they assume a single decision maker and therefore we have to look for the extension 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 at a satisfactory 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 in (Carlsson and Fullér, 1996; Lu *et al.*, 2007). Those approaches have been experienced in LitIO many times. Investigation of their suitability in the LitIO evaluation problem would require much investigations from other sciences, in particular management and psychology. For example, most

meetings are conducted online (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 to implement group decision making. Many references can be found at (Lu *et al.*, 2007; Rao, 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 can 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.

The linguistic scale presented in presented in Table 1 and Fig. 3 is not suitable for assigning weights to the decision makers or to the criteria. For that purpose we will use the importance scale presented in Table 2 (Lu *et al.*, 2007).

Table 2

Weights of a triangular distribution of a linguistic importance scale (Lu *et al.*, 2007)

Item of the Importance degrees scale	Numerical weights \tilde{A}
Absolutely unimportant	(0, 0, 1/6)
Unimportant	(0, 1/6, 1/3)
Less important	(1/6, 1/3, 1/2)
Important	(1/3, 1/2, 2/3)
More important	(1/2, 2/3, 5/6)
Strongly important	(2/3, 5/6, 1)
Absolutely important	(5/6, 1, 1)

9. Score Aggregation Method for Submission Ranking and Team Selection

We already concluded that the WSM approach best suits the submission ranking problem. All three analysed methods (WSM, WPM and Topsis) can be considered for team selection. However in this paper we focus on WSM, leaving application of the other two methods for a separate research.

We were looking for an extension of WSM to GDM, such that it would allow fuzzy numbers in the decision matrix, but 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 arrived at 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 (Chen *et al.*, 1992; Rao, 2007).

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 *et al.*, 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 and the team selection problem. Further we use the notations introduced in the third section.

The linguistic terms are converted to fuzzy numbers as it was previously described. 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 $\mu_{min}(x)$:

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

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

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

$$\mu_R(\tilde{A}) = \underset{x}{Sup}[\mu_{\tilde{A}}(x) \cap \mu_{min}(x)] \quad (7)$$

$$\mu_L(\tilde{A}) = \text{Sup}_x[\mu_{\tilde{A}}(x) \cap \mu_{max}(x)] \tag{8}$$

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

$$\mu_T(\tilde{A}) = (\mu_L(\tilde{A}) + 1 - \mu_R(\tilde{A}))/2 \tag{9}$$

Conversion of a fuzzy number to crisp value is illustrated in Fig. 4.

The values of the nine-item linguistic scale presented in Table 1 converted to crisp values are presented in Table 3. The values of the seven-item linguistic importance scale presented in Table 2 converted to crisp values are presented in Table 4. 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

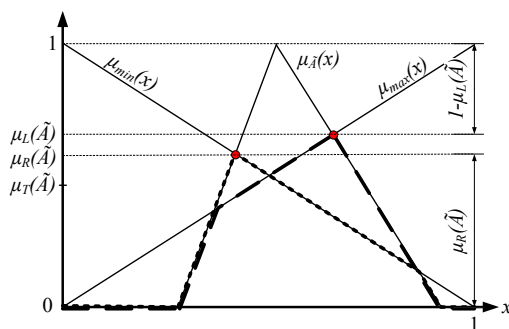


Fig. 4. Conversion of triangular fuzzy number to crisp value.

Table 3
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	0.92

Table 4
Calculating crisp values of the seven-item linguistic importance scale

The Importance Degrees	Fuzzy number \tilde{A}	$\mu_R(\tilde{A})$	$\mu_L(\tilde{A})$	$\mu_T(\tilde{A})$
Absolutely unimportant	(0, 0, 1/6)	1	0.14	0.07
Unimportant	(0, 1/6, 1/3)	0.86	0.29	0.21
Less important	(1/6, 1/3, 1/2)	0.71	0.43	0.36
Important	(1/3, 1/2, 2/3)	0.57	0.57	0.50
More important	(1/2, 2/3, 5/6)	0.43	0.71	0.64
Strongly important	(2/3, 5/6, 1)	0.29	0.86	0.79
Absolutely important	(5/6, 1, 1)	0.14	1	0.93

for the criterion 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.

The application slightly differs for the evaluation and for the team selection problems. We will start from the evaluation problem.

Let $A = \{A_1, A_2, \dots, A_n\}$ be a finite set of alternatives, $C = \{C_1, C_2, \dots, C_m\}$ be a finite set of criteria. Let $D = \{D_1, D_2, \dots, D_q\}$, $q \geq 2$ be a finite set of decision makers.

Each decision maker is assigned a linguistic weight of his/her importance and the linguistic value has been transformed to a fuzzy number and to its crisp value: $p = \{p_1, p_2, \dots, p_q\}$.

Each criterion is assigned a linguistic weight of its importance by each decision maker and transformed to a fuzzy number and then to its crisp value: $w_j = \{w_j^1, w_j^2, \dots, w_j^q\}$, ($j = 1, 2, \dots, m$).

Let $v_j^k(A_i)$ be the values of partial value functions of the performance of alternative A_i in terms of each criterion C_j by the decision maker D_k , where $i = 1, 2, \dots, n$, $j = 1, 2, \dots, m$, and $k = 1, 2, \dots, q$.

First the aggregated group weights for each criterion are calculated:

$$w_j = \frac{\sum_{k=1}^t w_j^k p_k}{\sum_{k=1}^t p_k}, j = 1, 2, \dots, m \quad (10)$$

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

$$v_j(A_i) = \frac{\sum_{k=1}^t v_j^k(A_i) p_k}{\sum_{k=1}^t p_k} \quad (11)$$

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

$$v(A_i) = \frac{\sum_{j=1}^m v_j(A_i)w_j}{\sum_{j=1}^m w_j} \quad (12)$$

Based on the calculated values, the ranking of the alternatives is performed. The above holds for submission evaluation problem.

For the team selection problem the values of the alternatives against each criterion are numeric, pre-calculated and given to the decision makers, i.e. $v_j(A_i) = x_{ij}$. x_{ij} was defined when defining the decision matrix 1). The total aggregated values for the team selection problem are calculated using this formula:

$$v(A_i) = \frac{\sum_{j=1}^m x_{ij}w_j}{\sum_{j=1}^m w_j} \quad (13)$$

Conclusions

In this paper we proposed to combine the group decision support algorithm combined with score aggregation method to be applied during evaluation and team selection process in the 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 of parts of it (i.e. those disclosed to the contestants) to the wider audience. Another important requirement was use of a 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 for the evaluation problem.

There is much more potential for the MCDA application for team selection problem. One reason is that the number of stakeholder is very small (from 5 to 12 contestants involved) and they are top students with good mathematical and algorithmical skills, which makes it easier to explain for them to accept mathematical decision making methods. Therefore several methods were chosen as possible for consideration for this problem. Before proposing to apply any of those methods in practice, the intermediate step would be to model the problem with data from previous years and to analyse differences between the models for the team selection problem.

Even though this paper presented the case of the Lithuanian Informatics Olympiad, the proposed solutions can be considered in other educational contexts as long as similar constraints are valid. The constraints include that the problem under consideration is a ranking, repeated, group decision making problem involving decision makers with a different level of their expertise skills as well as the need to present the decision making process to the stakeholders.

References

- Aruldoss, M., Lakshmi, T.M., Venkatesan, V.P. (2013). A survey on multi criteria decision making methods and its applications. *American Journal of Information Systems*, 1(1), 31–43.
- Belton, V. Stewart, T.J. (2003). *Multiple Criteria Decision Analysis: An Integrated Approach*. Boston, Kluwer Academic Publishing.
- 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. Berlin, Germany, Springer-Verlag.
- 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.
- Filipe, M., Ferreira, F., Santos, S. (2015). A multiple criteria information system for pedagogical evaluation and professional development of teachers. *Journal of the Operational Research Society*.
- French, S. (1988). *Decision Theory: An Introduction to the Mathematics of Rationality*. Chichester, Ellis Horwood.
- Ho, W., Dey, P.K., Higson, H.E. (2006). Multiple criteria decision making techniques in higher education. *International Journal of Educational Management*, 20(5), 319–337.
- Kahraman, C. (2008). *Fuzzy Multi Criteria Decision Making. Theory and Applications with Recent Developments* (vol. 16 of Optimization and its Applications). Springer.
- Keeney, R.L. Raiffa, H. (1976). *Decisions with Multiple Objectives: Preferences and Value Tradeoffs*. John Wiley & Sons.
- Kurilovas, E. Serikoviėnė, S. (2010). Learning content and software evaluation and personalisation problems. *Informatics in Education*, 9(1), 91–114.
- Kurilovas, E., Vinogradova, I., Serikoviėnė, S. (2011). Application of multiple criteria decision analysis and optimisation methods in evaluation of quality of learning objects. *International Journal of Online Pedagogy and Course Design*, 1(4), 62–76.
- 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).
- Oberti, P. (2004). D cision publique et recherche proc durale : illustration d'une d marche multicrit re   la localisation participative d'un parc  olien en r gion corse. In: *Actes des Journ es de l'Association Fran aise de Science Economique, Economie : aide   la d cision publique*. Universit  de Rennes.
<http://crereg.eco.univ-rennes1.fr/afse/TEXTES-PAR-SESS/A2/OBERTI.P.75.pdf>
- 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*. London, Addison-Wesley.
- Roy, B. (1996). *Multicriteria methodology for decision aiding*. Dordrecht, Kluwer Academic Publishers.
- 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, vol. 15). Scientific Committee of Lithuanian Informatics Olympiads. (2015).
- Shee, D., Wang, Y. (2008). Multi-criteria evaluation of the web-based e-learning system: a methodology based on learner satisfaction and its applications. *Computers and Education*, 50(3), 894–905.
- Simon, H.A. (1976). *Administrative Behavior*. New York, The Free Press.

- Siskosa, Y., Grigoroudis, E., Krassadakis, E., Matsatsinis, N. (2007). A multicriteria accreditation system for information technology skills and qualifications. *European Journal of Operational Research*, 182(2), 867–885.
- 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*. New York, Basel, Marcel Dekker.
- Triantaphyllou, E. (2000). *Multi-Criteria Decision Making Methods: a Comparative Study*. Kluwer Academic Publishers.
- Tsinidou, M., Gerogiannis, V., Fitsilis, P. (2010). Evaluation of the factors that determine quality in higher education: an empirical study. *Quality Assurance in Education*, 18(3), 227–244.
- Tzenga, G., Chiang, C., Lia, C. (2007). Evaluating intertwined effects in e-learning programs: a novel hybrid mcdm model based on factor analysis and dematel. *Expert Systems with Applications*, 32(4), 1028–1044.
- 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.
- Zeleny, M. (1982). *Multiple Criteria Decision Making*. New York, McGraw-Hill Book Company.
- Zhang, W. (2004). Handover decision using fuzzy MADM in heterogeneous networks. In: *Wireless Communications and Networking Conference, WCNC, IEEE* (vol. 2). 653–658.



dr. **J. Skūpienė** works at Informatics Methodology Department in Vilnius University Institute of Mathematics & Informatics. She has published about 20 scientific papers. She is a member of the Scientific Committee of National Olympiads in Informatics. She has been involved in many projects related to identifying, attracting and working with secondary and high school students gifted in IT. For a few years she was director of studies of Young Programmers School, since 2004 she has been a coordinator of Informatics division in the Lithuanian National Academy of Students. She is author/co-author of four books on algorithms and algorithmic problems.