INTERACTION OF CRITERIA IN GRADING PROCESS

Lavoslav Čaklović
Mathematical Department, University of Zagreb
Bijenička 30, Zagreb, Croatia
caklovic@math.hr
tel. +385-1-4605739

Abstract. The paper presents and discusses a grading problem in which additive utility is not suitable to express decision maker's preferences. We suggest two approaches how to solve it by using Potential Method. The first approach is subjective, whereas the other one is more objective and takes into account the full criteria profile of the students concerned.

The second approach introduces coalitions as new criteria and only those coalitions are considered that have strength over a certain threshold. This leads to a missing data problem with exact values. Finally, we present one idea often used in grading process, introducing a new criterion that takes into account oscillation of the the individual grades.

1. Introduction

1.1. **Importance of assessment process in education.** Assessment is one of the most important tasks in the whole teaching and learning process. Current education research show that multi-criteria assessment methods are widely used for evaluating not only students learning but for evaluating the university courses and the whole educational programs.

Determination of multi-criteria for assessing students and educational programs not only needs to meet the objectives of the students, but also should satisfy the requirements of the whole group. Each country has developed its own standards in educational assessment. The American National Council on Measurement in Education (NCME 1995) has adopted the *Code of Professional Responsibilities in Educational Measurement* to promote professionally responsible practice in educational measurement[1].

Responsibilities of various subjects in education arises, as it is written in the document, from either the professional standards of the field, general ethical principles, or both. This responsibilities are precised in the following list and we selected only the parts from those items from the list that concerns, directly or indirectly, assessment techniques.

Section 1: *Responsibilities of Those Who Develop Assessment Products and Services.* Those who develop assessment products and services, such as

2000 *Mathematics Subject Classification.* 90B50, 91B06.
Key words and phrases. Multicriteria decision making, interactive criteria, grading, coalition, Potential Method.

[1]Prepared by the NCME Ad Hoc Committee on the Development of a Code of Ethics.

classroom teachers and other assessment specialists, have a professional responsibility to strive to produce assessments that are of the highest quality. Persons who develop assessments have a professional responsibility to:

1.2 Develop assessment products and services that are as free as possible from bias due to characteristics irrelevant to the construct being measured, such as gender, ethnicity, race, socioeconomic status, disability, religion, age, or national origin.

1.3 Plan accommodations for groups of test takers with disabilities and other special needs when developing assessments.

1.11 Develop score reports and support materials that promote the understanding of assessment results

Section 2: *Responsibilities of Those Who Market and Sell Assessment Products and Services.* The marketing of assessment products and services, such as tests and other instruments, scoring services, test preparation services, consulting, and test interpretive services, should be based on information that is accurate, complete, and relevant to those considering their use. Persons who market and sell assessment products and services have a professional responsibility to:

2.5 Establish reasonable fees for assessment products and services.

2.10 Maintain a current understanding about assessment products and services and their appropriate uses in education.

Section 5: *Responsibilities of Those Who Score Assessments.*

The scoring of educational assessments should be conducted properly and efficiently so that the results are reported accurately and in a timely manner. Persons who score and prepare reports of assessments have a professional responsibility to:

5.1 Provide complete and accurate information to users about how the assessment is scored, such as the reporting schedule, scoring process to be used, rationale for the scoring approach, technical characteristics, quality control procedures, reporting formats, and the fees, if any, for these services.

5.2 Ensure the accuracy of the assessment results by conducting reasonable quality control procedures before, during, and after scoring.

5.3 Minimize the effect on scoring of factors irrelevant to the purposes of the assessment.

Section 8: *Responsibilities of Those Who Evaluate Educational Programs and Conduct Research on Assessments.* Conducting research on or about assessments or educational programs is a key activity in helping to improve the understanding and use of assessments and educational programs. Persons who engage in the evaluation of educational programs or conduct research on assessments have a professional responsibility to:

8.1 Conduct evaluation and research activities in an informed, objective, and fair manner.

8.8 Use multiple sources of relevant information in conducting evaluation and research activities whenever possible.

Another document which gives importance to assessment procedure in education is *Standards for Teacher Competence in Educational Assessment of Students*

(Sanders 1990.) developed by the American Federation of Teachers National Council on Measurement in Education National Education Association.

1.2. **Assessment techniques.** In a document *Teacher Preparation in California Standards of Quality and Effectiveness, Common Standards*[2] (Fortune 2003) the authors gives the criteria for evaluating the quality of educational programs. They are highly conscious of the importance of goals and subgoals in such evaluations and they state the explicit criteria for quality assessment. The standard technique in such procedures is to organize the subjects in decision process as the elements of the levels in an *hierarchical structure* in which the more important subject is placed in the higher level. The main *goal* is placed on the top of the structure followed by the less important goals, *attributes* or *criteria.* At the bottom level are objects that are evaluated, the *alternatives.*

The main step in multi-criteria decision making, including the classical multi-attribute utility (MAUT) model is aggregation. The aggregation procedure, roughly speaking, is done for each level with respect to criteria in upper level until we reach the bottom level. In MAUT, we aggregate one-dimensional utility functions U_i into a single global utility function and combine all the criteria. One of the simplest aggregation process is weighted arithmetic sum

$$U(x_1, \ldots, x_n) = \sum_{i=1}^{n} w_i U_i(x_i),$$

called additive utility. In grading process this is the most frequently used procedure.

In Potential Method, explained briefly in section 3 the principal object is preference graph while the aggregation procedure is done over the set of preference graphs for each criterion (attribute) rather than over the individual utility functions. This is more flexible approach which allows to treat the models with incomplete data without any additional transformation of the input data. At this stage the ELECTRE method and Potential Method have some points in common, although PM seems to be more sophisticated.

1.3. **Organization of the paper.** The intention of this article is to show how the PM can be adapted to serve as an evaluation tool in grading process. We focused our research at one of the most difficult problem in assessment procedure, the *interaction of criteria* where additive utility is not suitable to express decision maker's preferences. This procedure will be explained on the example that can serve as a model example in different circumstances.

The paper is organized as follows: first we state the problem in the form of an example. A short description of PM is given in section 3. In section 4 we explain two possibilities how PM can be applied to solve the problem. The first one is subjective, and another one takes into account the full criteria profile of the students under consideration. The second approach introduces the new criteria called coalitions and only the coalitions that have strength over the certain threshold are taken into consideration. This leads to a missing data problem with exact values solved by PM. Finally, we discuss an idea often used in grading process, which introduces a new criteria that considers oscillation of the individual grades.

[2]prepared by Committee on Accreditation Commission on Teacher Credentialing, State of California, June 1998, (Revised May 2002)

2. Dependence with respect to preferences

Attributes that decision maker uses to distinguish the alternatives are often linked by logical or factual interdependencies. In other words, they interact. A typical example can be found in a gastronomic context where the choice of wine depends upon the main dish. In some situations criteria have positive correlation, whereas in another, negative one. Bouyssou uses a notion of *concordance threshold* to express the strength of such coalitions, see (Bouyssou 1996, Ch. 6.4, p. 139).

2.1. **An example.** This example is from (Bouyssou 1996).

Example 2.1. Consider four students enrolled in an undergraduate programme consisting of three courses: Physics, Mathematics and Economics. For each course, a final grade between 0 and 20 is allocated. The results are given in Table 1. On the

	F	M	E
a	18	12	6
b	18	7	11
c	5	17	8
d	5	12	13

Table 1. Group profile

basis of these evaluations, it is felt that a should be ranked before b. Although a has low grade in Economics, he has reasonably good grades in both Mathematics and Physics which makes him a good candidate for an Engineering programme; b is weak in Mathematics and it seems difficult to recommend him for any programme with a strong formal component (Engineering or Economics). Using a similar type of reasoning, d appears to be a fair candidate for a programme in Economics. Student c has two low grades and it seems difficult to recommend him for a programme in Engineering or in Economics. Therefore d is ranked before c, and ranking is as follows:

$$a \succcurlyeq b \succcurlyeq d \succcurlyeq c.$$

Although these preferences appear reasonable, they are not compatible with the use of weighted average in order to aggregate the three grades. If we denote the weights (importance) of the courses by w_F, w_M i w_E it is evident that:

- ranking a before b implies that $w_M > w_E$:

$$18w_F + 12w_M + 6w_E > 18w_F + 7w_M + 11w_E \Rightarrow w_M > w_E,$$

- ranking d before c implies $w_E > w_M$:

$$5w_F + 12w_M + 13w_E > 5w_F + 17w_M + 8w_E \Rightarrow w_E > w_M,$$

which is contradictory.

In this example it seems that criteria interact. Such interactions, although not unfrequent, cannot be dealt with weighted mean value. They can be treated by fuzzy decision making approach, see (Grabish 1996). In the article we suggest a simple and flexible approach based on oriented graph approach called Potential

Method (PM), introduced for the first time in (Čaklović). Here is a brief description of PM.

3. Potential method, a brief description

3.1. **Single criterion.** For simplicity, let us suppose that one decision maker makes pairwise comparisons of alternatives, the set of alternatives denoted by V, and uses a single criterion. We give a brief description of Potential Method, details can be found in (Čaklović).

A pair $\alpha = (u, v) \in V \times V$ is declared to be an arc of a directed graph if v is more preferred than u. An un-compared pair is not adjacent in the graph. The set of all arcs is denoted by $\mathcal{A}$. A function $F : \mathcal{A} \to \mathbb{R}$ which assigns to each arc $\alpha \in \mathcal{A}$ its weight of preference is called a `preference flow`. The flow component $F(\alpha)$ is usually taken from a given interval; here we use interval $[0, 4]$ of real numbers. Evidently, preference flow is always non-negative and can be represented as an $m \times 1$ matrix. Oriented graph $(V, \mathcal{A})$ is called a `preference graph`. The preference graph is `complete` if each pair of alternatives is compared i.e. if for each pair (i, j) of vertices $(i, j) \in \mathcal{A}$ or $(j, i) \in \mathcal{A}$. For a given flow F on the preference graph $(V, \mathcal{A})$ and $\alpha \in \mathcal{A}$ we use a convention $F(-\alpha) := -F(\alpha)$.

Let us denote by n and m the cardinality of V and $\mathcal{A}$ respectively. Incidence matrix of the preference graph is denoted by B and it is $m \times n$ matrix defined by

$$B_{\alpha,i} = \begin{cases} -1, & \text{if } \alpha \text{ leaves } i \\ 1, & \text{if } \alpha \text{ enters } i \\ 0, & \text{otherwise.} \end{cases}$$

Let F be a given preference flow, and B incidence matrix of the graph. Moreover, let us assume that the graph is weakly connected (connected in the sequel). System

$$B^\tau BX = B^\tau F, \quad \textstyle\sum_{i=1}^{m} X_i = 0 \tag{3.1}$$

has a unique solution called `normal integral` of F. One can think of the first equation in (3.1) as the normal equation associated to $BX = F$. The potential difference BX of normal integral is the best approximation of F by column space of incidence matrix. Sometimes, a function $X : V \to \mathbb{R}$ is called potential and that is the reason why the method get its name: *Potential Method*. One can think of X as utility function. If the graph is not connected, the normal integral is unique on each connected component of the graph. To obtain a ranking, after having X, the following formula can be used

$$w = \frac{a^X}{\|a^X\|_1}, \quad a > 0$$

where exponent function of X is defined componentwise, i.e. $(a^X)_i = a^{X_i}$, and $\|\cdot\|_1$ is l_1 norm. Parameter a can be arbitrary, currently we use value $a = 2$.

`Measure of inconsistency` is defined as

$$\text{Inc}(F) = \frac{\|F - BX\|_2}{\|BX\|_2},$$

where $\|\cdot\|_2$ denotes 2-norm and $\beta = \arctan(\text{Inc}(F))$ is an `angle of inconsistency`. Ranking is considered 'acceptable' if β is less than 12 degrees. The last statement should not be taken for granted, as there is no serious argument to support it due to the freshness of the method.

3.2. **Consensus flow.** If more than one criterion is present then, each criterion C_i generates its own graph $(V, \mathcal{A}_i)$ and its own flow F_i. Let us denote the weight of the i-th criterion by w_i. We are going to describe a procedure of making a `consensus graph` $(V, \mathcal{A})$ and `consensus flow` F for the group of all criteria.

First, for a given pair $\alpha = (u, v)$ we calculate

$$F_\alpha := \sum_{\substack{i=1 \\ \pm\alpha\in\mathcal{A}_i}}^{k} w_i F_i(\alpha) \tag{3.2}$$

where the term $w_i F_i(\alpha)$ contributes if and only if $\pm\alpha \in \mathcal{A}_i$ i.e. if and only if $F_i(\alpha)$ or $F_i(-\alpha)$ is defined. If this sum is non-negative, then we put α in the set of arcs $\mathcal{A}$ and $F(\alpha) := F_\alpha$. Otherwise, we define $-\alpha = (v, u)$ as an arc in $\mathcal{A}$ and $F(-\alpha) := -F_\alpha$. The flow F becomes a non-negative flow that is called *consensus flow*. It can happen that consensus graph has a cycle. Anyway, normal integral of F exists and it is unique. The presence of cycles can only generate bigger inconsistency $\mathrm{Inc}(F)$.

3.3. **Consensus flow for decision table.** For decision table (such as Table 2) the consensus flow, defined on the graph with actions $a_1, \ldots a_m$ as vertices and with states $\theta_1, \ldots \theta_n$ as criteria, according to formula (3.2) is defined by

$$F_{jk} = \sum_i P(\theta_i)(v_{ki} - v_{ji}), \quad k, j = 1, \ldots, m. \tag{3.3}$$

Here F_{jk} denotes the flow component on arc (j, k) and represents the kj component of the `flow matrix`. Using formula (3.1) normal integral X can be calculated. Note that the above flow is complete and the flow matrix exists only for complete flows. On the other side, classical utility theory assigns to each action a_i its utility

$$U(a_i) := \textstyle\sum_j P(\theta_j) v_{ij}. \tag{3.4}$$

The following theorem gives motivation for definition (3.3) by proving equivalence of PM and the expected utility. The proof can be found in (Čaklović and Šego 2002).

Theorem 3.1. *Ranking over the set of alternatives given by expected utility is the same as the ranking given by Potential Method. More precisely, for each* $k = 1, \ldots, n$

$$X_k = U(a_k).$$

		States of nature			
		θ_1	θ_2	$\cdots$	θ_n
Actions	a_1	v_{11}	v_{12}	$\cdots$	v_{1n}
	a_2	v_{21}	v_{21}	$\cdots$	v_{21}
	.	.	.	$\cdots$	.
	a_m	v_{m1}	v_{m2}	$\cdots$	v_{mn}

TABLE 2. Decision table.

It is evident that the normal integral, expressed in terms of the flow-matrix, is given by

$$x_i = \frac{1}{n}\sum_{j=1}^{n} F_{ij}, \tag{3.5}$$

i e. X is an arithmetic mean of columns in the flow-matrix.

4. Analysis of example 2.1 by Potential Method

We shall explain some ideas how to model the situation in example 2.1 using the Potential Method.

4.1. **Subjective approach.**

4.1.1. *One-level model.* Following the reasoning of decision maker in the example we obtain the following preference graph on the set of students profiles. in which

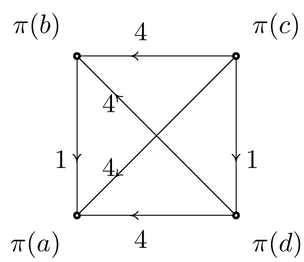

Figure 1. Subjective preferences.

the preferences of $\pi(a)$ and $\pi(b)$ compared to $\pi(c)$ and $\pi(d)$ are the same. Also, $\pi(d)$ has a small preference against $\pi(c)$ and the same preference has $\pi(a)$ against $\pi(b)$. The preference flow matrix (F_{ij}) and the normal integral are given in Table 3. Weak preference relation generated by X is $a \succcurlyeq b \succcurlyeq d \succcurlyeq c$, i. e. the same one as suggested in example 2.1. Calculation was done by WEB interface on URL

	a	b	c	d	X
a	0	1	4	4	2.25
b	−1	0	4	4	1.75
c	−4	−4	0	−1	−2.25
d	−4	−4	1	0	−1.75

Table 3. Flow matrix.

`http://decision.math.hr/Self/`. The snapshot of the interface is given bellow in Table 4 as well as the results obtained after the processing data, Table 5.

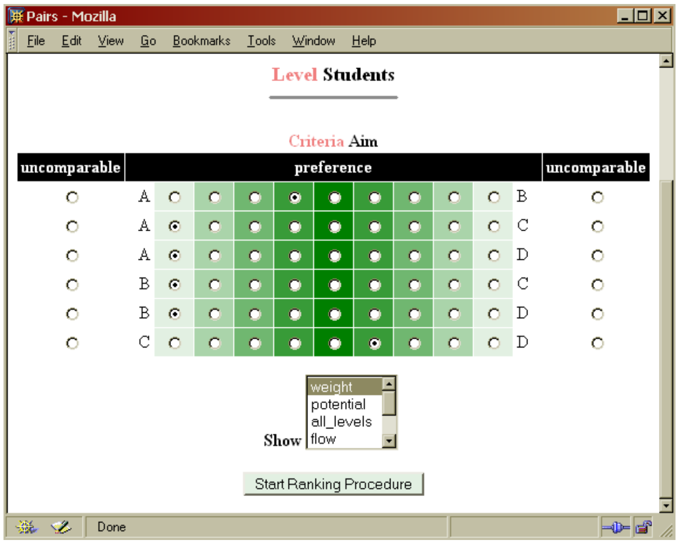

TABLE 4. WEB interface for PM.

```
Level: students ( Norm = 1.000 )
Comp_1 Weight = 1.000 InvInc = 0.124 ( Angle = 7.07 deg )
    Nodes:
          a    0.551    ( X = 2.25 )
          b    0.390    ( X = 1.75 )
          c    0.024    ( X = -2.25 )
          d    0.034    ( X = -1.75 )

     Total weight = 1.000
```

TABLE 5. Obtained result.

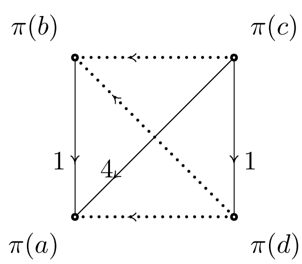

FIGURE 2. Incomplete preference graph.

4.1.2. *Missing data.* In this model the corresponding preference graph is incomplete and there is no explicit formula for the solution of equation (3.1). Preference graph is given in Figure 2. Dotted lines represent 'forgotten' comparisons from the previous case of complete graph. In Table 1 there are two classes of students. One class with 36 points and another one with 30 points. Students a and c can be considered as the representatives of their class. Preference $\pi(c) \xrightarrow{4} \pi(a)$ serves as the preference between those classes while preferences $\pi(b) \xrightarrow{1} \pi(a)$ and $\pi(c) \xrightarrow{1} \pi(d)$ describe the relationship within each class. Normal integral and weights obtained from this preference graph are given in Table 6

```
Level: students ( Norm = 1.000 )
Comp_1 Weight = 1.000 InvInc = 0.000 ( Angle = 0.00 deg )
   Nodes:
         a   0.593   ( X = 2.00 )
         b   0.296   ( X = 1.00 )
         c   0.037   ( X = -2.00 )
         d   0.074   ( X = -1.00 )
    Total weight = 1.000
```

Table 6. Forgotten data.

4.1.3. *Hierarchical model.* There are two groups of students, group $AB = \{a, b\}$ and group $CD = \{c, d\}$. The following construction leads to the same rating as before. It is reasonable to compare the groups first, and after that the students within each group. The group AB has preference over the group CD. Within the group AB student a has a preference against student b, and within the group CD student d has preference against student c. Using the same WEB interface we obtain the result given in Table 7. Let us point out that there are two connected components

```
Level: students ( Norm = 1.000 )
Comp_1 Weight=0.667 InvInc = 0.000 ( Angle = 0.00 deg )
    Nodes:
          a   0.409   ( X = 0.33 )
          b   0.258   ( X = -0.33 )
Comp_2 Weight=0.333 InvInc = 0.000 ( Angle = 0.00 deg )
    Nodes:
          c   0.147   ( X = -0.17 )
          d   0.186   ( X = 0.17 )

     Total weight = 1.000
```

Table 7. Hierarchical model.

in the level 'students' with weights 0.667 and 0.333 respectively. Students' rating is the same as in one-level model.

4.2. **Exact data approach.** In this approach we use the data from the given profile, Table 1 page 4. Direct calculation of flow matrix and normal integral, as it was pointed out in Theorem 3.1, is equivalent to standard average utility approach and weighted arithmetic mean. According to formula (3.3), the component of the flow-matrix at place ab, for example, equals

$$F_{ab} = (18 - 18) + (-7 - 12) + (11 - 6) = 0 - 5 + 5 = 0,$$

and at place ac

$$F_{ac} = 18 - 5 + (12 - 17) + (6 - 8) = 13 - 5 - 2 = 6.$$

The flow-matrix is given in Table 8. The last column in the table is normal integral calculated as arithmetic mean of columns of the flow-matrix, according to formula (3.5). Column Av, in the table, gives the average note of each student. The conclusion is that potential X, as well as average note can't distinguish students a and b, and c and d. To solve the problem we have to take into account interaction

	a	b	c	d	X	Av
a	0	0	6	6	3	9
b	0	0	6	6	3	9
c	−6	−6	0	0	−3	7.5
d	−6	−6	0	0	−3	7.5

TABLE 8. Flow-matrix of the group profile.

among criteria and recalculate the normal integral.

4.3. **Adding new criteria.**

4.3.1. *Coalition as criterion.* The main cause of non-additivity of criteria is their interaction called `coalition`. It seems reasonable that coalition of two criteria become a new criterion. We shall introduce three coalitions: FM, ME, FE made of two criteria each. The `strength` of coalition is the sum of points of each criterion in coalition. We say that coalition is `weak` if its strength is less than 20 points, otherwise it is `strong`[3]. We shall take into account only those coalition that are

	F	M	E	FM	ME	FE
	5	5	5	3	3	3
a	18	12	6	30	$*$	24
b	18	7	11	25	$*$	29
c	5	17	8	22	25	$*$
d	5	12	13	$*$	25	$*$

TABLE 9. Coalition of criteria.

strong. The new table is Table 9, where $(*)$ denotes weak coalitions which are not taken into account. One can think of that table as the table with missing data.

[3]This threshold of 20 points seem to be reasonable because the maximum grade for each criterion is 20.

In the second row we specified the relative weight of each criterion including the coalition. We may point out that single criteria (F, M, E) have relative weight of 5 and the relative weight of each coalition is 3. We expect that coalition should generate a small perturbation of an influence that comes from main criteria[4]. Obtained ranking is

	rank
a	0.397
b	0.366
d	0.129
c	0.108

which is acceptable from decision maker's point of view in example 2.1.

4.3.2. *Oscillation as criterion.* Some teachers use `oscillation` as an extra criterion, where oscillation of a student is defined as the range of obtained points. A student is better if its oscillation is lower[5]. In this case decision table is

	F	M	E	Osc
	4	4	4	1
a	18	12	6	12
b	18	7	11	11
c	5	17	8	12
d	5	12	13	8

where Osc denotes oscillation. In the second row we specified the relative weight od main criteria and oscillation. Calculated weights are:

	rank
b	0.298
a	0.297
d	0.204
c	0.200

In this model b is still better than a. It seems that introducing oscillation, as a new criterion, is not good enough to explain the interaction of criteria in example 2.1.

[4] At this point we should fix one parameter more called `flow-norm`, and its value is 2. Definition of flow–norm and its influence on ranking procedure is explained in (Čaklović and Šego 2002).

[5] This statement should be reconsidered in each situation and its acceptance depend upon the situation and decision context.

5. Conclusion

Potential Method was primary developed to treat subjective data in the same manner as Saaty's eigenvalue method does, (Saaty 1996). The interface for collecting data can be the same for both methods but they are different in the core i. e. in numerical treatment of the data.

Decision support based on PM has URL `http://decision.math.hr/` and it's development was supported my the Ministry of Science and Technology of Republic of Croatia. Group decision support is not available at the moment, because the site is under the severe reconstruction. In this article we described only those features of PM that are available on the site at this moment. The interface for exact data and group decision support are in the testing phase and will be available in the future.

Another project of interest for education management based on PM is the scheduling method, some kind of weighted topological sorting, developed as a tool in creation of new university programmes. The tool is based on the criteria-alternatives duality in self ranking procedure in group decision.

Let us discuss a bit more the responsibilities described in *Code of Professional Responsibilities in Educational Measurement* and a possible role of PM in educational process. In each decision method based on criteria–alternatives hierarchical relationship it is easy to see if the assessment satisfy the ethnical requirement 1.2 and 1.3 because the criteria should be stated explicitly. The same is true for the requirement 5.3.

Graph interpretation of decision maker's preferences is simple and understandable for both; for evaluator (teacher) and for evaluating object (student, educational programm developer). Even the students and educational programm developers can be engaged in the evaluating process. This means that requirements 5.1, 5.2, 8.1, 8.2 are easy to satisfy. Let us just mention that fuzzy approach to decision making, although widely used, is hardly understandable to the majority of potential users.

References

Bouyssou, D. at al. 1996. *Evaluations and decision models, A critical perspective*, Kluwer: International series in operational research & management science

Cynthia B. Schmeiser at al. 1995, *Code of Professional Responsibilities in Educational Measurement*, The American National Council on Measurement in Education (NCME), `http://www.natd.org/Code_of_Professional_Responsibilities.html/`.

Čaklović L. *Decision Making by Potential Method*, preprint

Čaklović, L. and Šego, V. 2002. Potential Method applied on exact data, *Proceedings of* 9^{th} *Int. Conf. on Oper. Res. KOI2002*, Croatian Operational Research Society, 237–248. Available from: `http://www.math.hr/~caklovic/Decision/Download/exact.pdf`

Fortune, M. at al. 2003. *Standards Of Quality and Effectiveness for Preliminary Administrative Services Credentials*, Handbook for Teacher Educators and Accreditation Team Members, California Commission on Teacher Credentialing, Sacramento, California `http://www.ctc.ca.gov/educator-standards/SVC-Preliminary-Admin.pdf`.

Grabish, M. 1996. The applications of fuzzy integrals to multicriteria decision making, *European Journal of Operational Research* **89**:445–456

Saaty, T. L. 1996. *The Analytic Hierarchy Process*, RWS Publications, Pittsburg

Sanders, James R. at al. 1990. *Standards for Teacher Competence in Educational Assessment of Students*, American Federation of Teachers National Council on Measurement in Education National Education Association, http://www.unl.edu/buros/article3.html.

PMF–Mathematical Department, Bijenička 30, 10 000 Zagreb, Croatia
E-mail address: caklovic@math.hr