Fuzzy measure identification for multicriteria decision making

Milica Savić

Contents — **In this paper, we address the problem of identifying the hard-to-determine fuzzy measures for multicriteria decision making problem with interacting criteria. We devise approach to the problem, based on M. Grabish's fuzzy measure identification algorithm. Fuzzy measure identifying procedure, based on needs of specified task, is described and example is shown.**

Keywords—**Criteria interaction, Fuzzy measure identification, Multicriteria decision making.**

I. Introduction

THE main appeal of using fuzzy measures in multicriteria decision making is their ability to model relative importance of decision making criteria and their complex interactions. Main difficulty, concerning their practical use, is the necessity of defining 2^N coefficients for N criteria problem in order to define a fuzzy measure. For most applications, experts are intended to assess coefficients, but this is still very limiting, since for the large number of criteria, the task of assessing coefficients becomes too big and difficult for any kind of practical use. One approach to the problem is described in this paper, as well as the application on the considered ARDS (Acute Respiratory Distress Syndrome) classification problem. ARDS classification problem is a multicriteria decision making problem where doctor is asked to classify patient in an intensive care unit as being in phase N, I, II, III or IV (most severe phase) of a respiratory failure, based on 5 given symptoms.

II. Multicriteria fuzzy decision making

The known general model of a multicriteria decision making system is shown in Fig. 1.

Milica Savić, Računarski fakultet, Srbija (e-mail: msavic03@raf.edu.rs).

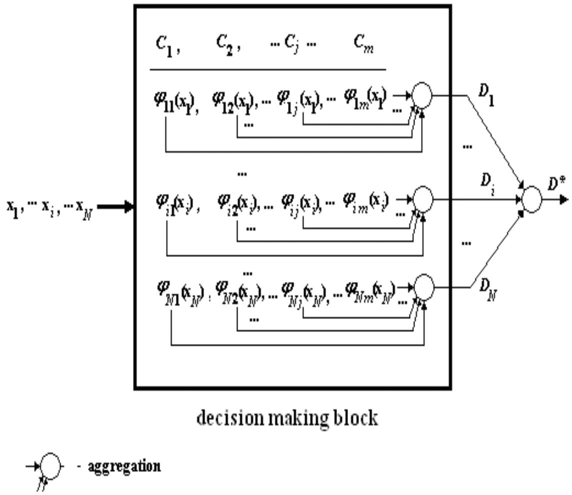

Fig. 1. Fuzzy multicriteria decision-making system.

In Fig. 1: $\mathbf{x}_i$, i = 1,2, ... , N, are vectors of object properties, which are considered in decision making process; C_j, j = 1,2, ... , m, are decision making criteria; $\varphi_{ij}(\mathbf{x}_i)$, i = 1,2, ..., N, j = 1,2, ... ,m, are scores, i.e. degrees in which an object $\mathbf{x}_i$ (or its property) satisfies the criteria C_j. D_i, i = 1, 2, ... , N, are decisions (performance indices) of an object $\mathbf{x}_i$ with respect to all the criteria C_j. Decisions D_i are obtained by aggregation of information $\varphi_{ij}(\mathbf{x}_i)$, using suitable aggregation operation. The decision D^*, on object $\mathbf{x}_i$ that best satisfies all criteria C_j, j = 1, 2, ... , m, is obtained by aggregation of decisions D_i, - using suitable aggregation operation, appropriate for the considered problem. In fuzzy multicriteria decision-making systems, a score has the following property: $\varphi_{ij}(\mathbf{x}_i) = \mu_{ij}(\mathbf{x}_i) \in [0,1]$ and is treated as a fuzzy measure. Thus, the aggregation process of fuzzy information is an important element of fuzzy decision making system.

Some of the most commonly used aggregation operators are: family of quasi-arithmetic means operators (such as simple arithmetic mean, geometric, harmonic means, etc.), median (taking into account not the values themselves but only their ordering), weighted minimum, weighted maximum, ordered weighted averaging operators (OWA). All these operators are idempotent, continuous, and monotonically non decreasing (ranging between min and max). Their main common characteristic is that they all are averaging operators. Reader interested in more details, analysis, comparisons and classifications of family of operators can consult references [1], [2], [3].

All these operators have some drawbacks: Some do not possess all the desirable properties (e.g. quasi-arithmetic means are not stable under positive linear transformation), and some seem to be too restrictive (arithmetic sums, OWA, etc.). The main point here is that no one is able to model interaction between criteria in some understandable way.

Sugeno proposed the concept of non-additive fuzzy measure and fuzzy integral in 1974, [4]. For interacting criteria decision making, the Choquet integral represents a suitable aggregation operator.

III. Fuzzy Measure and the Choquet Integral

A fuzzy measure (or the Choquet capacity) on $C = \{C_1, ..., C_m\}$ is a monotonic set function μ: $P(C) \to [0,1]$, where $P(C)$ is the power set of the set C, with $\mu(\emptyset)=0$ and $\mu(C)=1$. Monotonicity means that $\mu(S) \le \mu(T)$, whenever $S \subseteq T \subseteq C$. An interpretation of $\mu(S)$ can be that it is the weight related to the subset S of criteria.

Given μ, the Choquet integral of $x \in (\mathbf{R}^+)^n$ with respect to μ is defined by

$$Ch_\mu(x) := \sum_{i=1}^{n} (x_{(i)} - x_{(i-1)}) \mu(\{(i),...,(n)\}). \tag{1}$$

In (1) $(\cdot)$ means a permutation of the elements of C such that $x_{(1)} \le ... \le x_{(N)}$ and $x_{(0)} = 0$.

A more exhaustive study of the Choquet integral properties can be found in [5], but will not be further discussed here.

IV. The Choquet Integral for Interacting Criteria Modeling

As an illustrative example for interacting criteria decision making, the student ranking problem is commonly used [5]: Students are evaluated according to their level in 3 subjects: mathematics, physics and literature. More importance is attributed to mathematics and physics, and the two are considered equally important. Coefficients of importance are chosen accordingly: 3 for math, 3 for physics, and 2 for literature.

Computing the average evaluation of the students by using a simple weighted mean, and with marks given on scale from 0 to 20, 3 students are evaluated in Table 1.

TABLE 1: WEIGHTED MEAN STUDENT EVALUATION.

Student	Mathematics	Physics	Literature	Global evaluation (weighted mean)
A	18	16	10	15.25
B	10	12	18	12.75
C	14	15	15	14.62

The shown weighted mean student ranking is not satisfactory if the school ranking them wants to favor students without weak points. In the shown ranking, student A has severe weakness in literature, but is still ranked higher than student C, which has no weak points. This is due to too much importance being given to mathematics and physics, which are in a sense redundant, since usually, students good at mathematics are also good at physics (and vice versa). This kind of evaluation tends to overestimate (resp. underestimate) students good (resp. bad) at mathematics and/or physics. Through use of the Choquet integral, a more complex decision making process reflecting criteria interaction can be modeled.

For the student ranking example, suppose the decision makers preferences are:

1. Scientific subjects (math, physics) are more important.
2. Scientific subjects are more or less similar, and students good at mathematics (resp. physics) are in general also good at physics (resp. math), so that students good at both must not be too favored.
3. Students good at mathematics (or physics) *and* literature are rather uncommon and must be favored.

These can be directly translated in term of fuzzy measure as:

1. μ ({mathematics}) = μ({physics}) = 0.45,
   μ({literature)} = 0.3
   (relative importance of scientific versus literary subjects)

2. μ({mathematics, physics}) = 0.5 < μ({mathematics}) + μ({physics})
   (redundancy between mathematics and physics)

3. μ({mathematics, literature}) = μ({physics, literature}) = 0.9 > 0.45 +

0.3 (support between literature and scientific subjects)

The idea is that superadditivity of the fuzzy measure implies synergy between criteria, and subadditivity implies redundancy. Note that it is up to expert to scale these values to the extent that he feels expresses the importance and interaction.

Applying the Choquet integral with the above fuzzy measure leads to the following new global evaluation shown in Table 2:

TABLE 2: THE CHOQUET INTEGRAL STUDENT EVALUATION.

Student	Mathematics	Physics	Literature	Global evaluation (the Choquet integral)
A	18	16	10	13.9
B	10	12	18	13.6
C	14	15	15	14.9

Here, students are properly ranked in accordance to the preference relation.

V. FUZZY MEASURE IDENTIFICATION PROBLEM

The use of fuzzy measures is increasingly difficult with the increase of the number of criteria, since for the N criteria decision problem, one has to identify 2^N coefficients in order to define a fuzzy measure. Often, presence of an expert is assumed for the assessment of coefficients. Even so, the assessment step is still considered difficult. Especially tricky part is assessment of interaction of three and more criteria taken together. In [6], where an expert is asked to input the 2^5 coefficients for determining the severity of respiratory distress, this task alone makes the model an impractical one, even for experts. In [6], developed Java application is described, and here we will not go into details, but will use it to illustrate the difficulty of identifying interactions amongst criteria.

On Fig. 2, the required coefficients for 5 criteria (symptoms of illness) are shown. The developed application relaxes the difficulty of the task on expert by computing the default coefficients for the non-interacting criteria, so the expert using the application is prompted to alter them according to their synergy or redundancy. What makes the task additionally difficult is the fact that the coefficients of the fuzzy measure must obey monotonicity constraints. Application aids the user by warning him when monotonicity is lost, and points out the conflicted coefficients. This is still a tedious task, and if the number of classifying criteria used in application was to grow, rather than improving the model, this could lead to an unusable application instead – that is, the task of scaling alone would be too much to grasp.

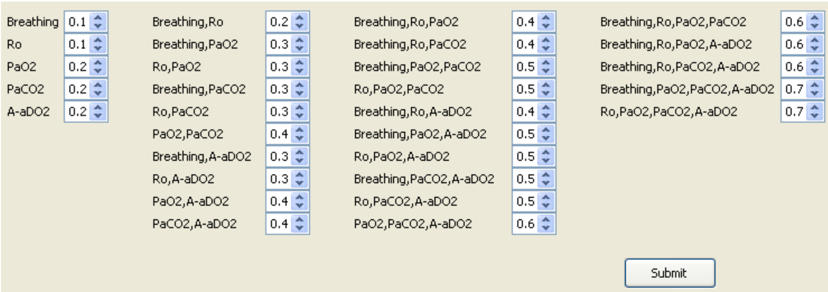

Fig. 2. Coefficients for 5 criteria.

To address the problem, a suitable algorithm could be used to assess the coefficients not altered by the expert using the application, while the coefficients altered by the expert would remain untouched by the algorithm. The next section describes M. Grabish's algorithm proposed in [7] for online fuzzy measure learning based on given learning datum. Despite the fact that in this paper we are not treating ARDS classification problem as an online learning problem, we propose to utilize one key aspect of M. Grabish's algorithm which differentiates this algorithm from several other algorithms devised for the same purpose (fuzzy measure identification), for instance, one proposed by Mori and Murofushi in [8].

VI. ALGORITHM FOR IDENTIFYING FUZZY MEASURES

Algorithm proposed in [7] arranges the coefficients for which the learning data are lacking to get the coefficients as homogeneous as possible, meaning distance from neighbor nodes should be as equal as possible. It consists of 2 basic steps best explained using lattice representation of the coefficients shown in Fig. 3.

For the description of the algorithm, the following terminology is used:

The lattice of a fuzzy measure is made from nodes related by links. The lattice has $n + 1$ horizontal layers, numbered from 0 (for the layer containing only $\mu_{\emptyset}$) to n (for the layer containing only μ_X). A path is a set of chained links, starting from the node $\mu_{\emptyset}$ and arriving to the node μ_X. For a given node in layer l, its lower neighbors (resp. upper neighbors) are the set of nodes in the layer l-1 (resp. $l + 1$) linked to it. There are l lower neighbors and $n - l$ upper neighbors.

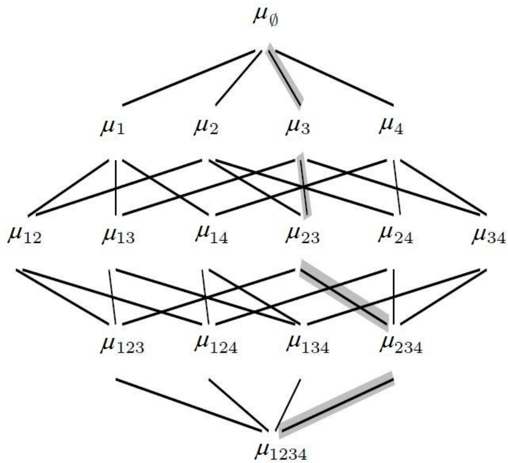

Fig. 3. Lattice of the coefficients of a fuzzy measure (n=4) with one path highlighted

The author of the algorithm describes it in 2 steps as follows:

step 0: The fuzzy measure is initialized at the equilibrium state (all criteria of equal importance, no interaction).

step 1: For a given learning datum x, we modify only the coefficients on the path involved by x in order to decrease the error, as in a gradient descent algorithm. The modification is done in order to preserve the monotonicity property *on the path*. Also, monotonicity is checked for neighboring nodes, but only for the nodes already modified in previous steps. This is done for all learning data, several times.

step 2: If there are too few learning data, then some nodes may have been left unmodified. These nodes are modified here in order to have the most equilibrated lattice, i.e. distance from neighbors should be as equal as possible. Multiple iterations of this step are also reasonable.

The algorithm is used for pattern recognition and classification based on fuzzy integral in [7], [9], and the crucial point is the identification of the fuzzy measures using training data. In [7], adaptation of the above described algorithm is named Heuristic Least Mean Squares (HLMS). Principle is the same, and the alteration is in coefficient adaptation: for the targeted class of the training datum, coefficients are increased, and for all other decreased. HLMS algorithm is tested in [7] in terms of speed and convergence on test data: cancer data (284 training examples with 9 attributes for 2 classes) and simulated data (200 training examples with 5 attributes for 2 classes). The algorithm is compared to constrained least mean squares algorithm (CLMS), and has superior classification rate (77.4% for cancer data) in comparison to CLMS (72.9% for cancer data). Also, the CPU time for tested cancer data is significantly smaller for the proposed Heuristic least mean squares algorithm, and there is a slight loss in optimality.

ARDS classification based entirely on HLMS with no expert present, but relying on training data alone, is also feasible. Approach described in this paper is a step in that direction, providing experts an aid for complex multicriteria ARDS classification thru utilization of specific lattice structure of fuzzy coefficients.

For a more detailed description and analysis of HLMS algorithm, reader is referred to [7].

VII. ARDS CLASSIFICATION PROBLEM

Membership degrees for each phase of acute respiratory distress syndrome, for the given numerical values of symptoms, indicated by Table 3, are determined by using trapezoidal membership function, [10] and [11], and given by Table 4. In Table 3, in column “Phase”, N is normal condition of a patient, I – is first (least severe) phase of respiratory distress (injury and resuscitation), II – the second phase of respiratory distress (subclinical), III – the third phase (established respiratory distress), and IV – the fourth phase of distress (severe respiratory failure). The features *Breathing* and *Rö* are expressed verbally, and given subjective membership degree (fuzzy sets theory). Other features are characterized by approximate intervals of numerical values. For these features to be interpreted as fuzzy sets ‘x is approximately in the interval $[b, c]$’, they must be characterized by an order quadruple $A = (a, b, c, d)$, fuzzy trapezoidal number, [10]. Characteristic values of the criteria for determining the severity of respiratory distress, given by Table 3, are represented by fuzzy intervals formed on basis of experience, and given by Table 4.

TABLE 3: DECISION-MAKING PARAMETERS.

Phase	Breathing	Rö	PaO2	PaCO2	A-aDO2
N	-	-	80 – 100	35 – 45	5 – 10
I	normal	no changes	70 – 90	30 – 40	20 – 40
II	mild to moderate tachypnea	minimal infiltrates	60 – 80	25 – 35	30 – 50
III	increasing tachypnea	confluence of infiltrates	50 – 60	20 – 35	40 – 60
IV	obvious respiratory failure	generalized infiltrates	35 – 55	40 – 55	50 – 80

TABLE 4: FUZZY DECISION PARAMETERS.

Phase	PaO2	PaCO2	A-aDO2
N	(70,80,100,110)	(30,35,45,50)	(0,5,10,15)
I	(50,70,90,110)	(25,30,40,45)	(10,20,40,50)
II	(40,60,80,100)	(20,25,35,40)	(20,30,50,60)
III	(40,50,60,70)	(10,20,35,45)	(30,40,60,70)
IV	(30,35,55,60)	(30,40,55,65)	(40,50,80,90)

For a patient with symptoms described in Table 5, the decision making table is given (Table 6).

TABLE 5: SYMPTOMS OF A PATIENT TO BE DIAGNOSED.

Breathing	Rö	PaO2	PaCO2	A-aDO2
moderate tachypnea	confluence of infiltrates	50	32	31

TABLE 6: DECISION-MAKING TABLE FOR A PATIENT.

Phase	Breathing	Ro	PaO2	PaCO2	A-aDO2
N	0	0	0	0.4	0
I	0	0	0	1	1
II	0.9	0	0.5	1	1
III	0.3	0.8	1	1	0.1

Suppose physician's preferences for the given symptoms are: "Features *Breathing* and *Rö* are less important than other and between features PaO_2 and $PaCO_2$ there exists synergy". Expressed by a fuzzy measure these preferences could be:

For the first preference:

μ(*Breathing*)= μ(*Rö*)=0.1

$\mu(PaO_2)$= $\mu(PaCO_2)$= $\mu(A\text{-}aDO_2)$=0.2.

For the second preference:

$\mu(PaO_2,PaCO_2)$=0.5> $\mu(PaO_2)+\mu(PaCO_2)$=0.4.

To obtain evaluation for each phase of the illness using the Choquet integral, the following indexes of importance need to be defined: $\mu(PaCO_2, A\text{-}aDO_2)$, $\mu(Breathing, PaCO_2, A\text{-}aDO_2)$, $\mu(Breathing, PaO_2, PaCO_2, A\text{-}aDO_2)$, μ(*Rö*, PaO_2, $PaCO_2$), μ(*Breathing, Rö*, PaO_2, $PaCO_2$). For the given example, with a slight synergy between criteria PaO_2 and $PaCO_2$, and no other interactions, the decision makers reasoning for assessing other indexes where the two interacting criteria appear together could be to simply propagate the synergy (wherever the two interacting symptoms appear together, increase the index of importance, so for instance $\mu(Breathing, PaO_2, PaCO_2, A\text{-}aDO_2)$ would be 0.8). This approach, let's call it propagated interaction approach, would only be applicable in cases like this, where monotonicity is not lost by this kind of reasoning. If the synergy between 2 symptoms was greater, for instance $\mu(PaO_2, PaCO_2)$=0.8, we would have lost monotonicity, and $\mu(Breathing, PaO_2, PaCO_2, A\text{-}aDO_2)$ would be 1.1, which is greater than 1 (index of importance of all indexes taken together), so the decision maker would resort to scaling the conflicted values. This does not seem practical, especially for large number of criteria. Different approach comes from reasoning used in HLMS algorithm – for the lacking information, try to get indexes as homogenous as possible. Essentially, we use the second step of the algorithm for the hard-to-assess indexes, while indexes of importance identified by the expert decision maker are treated as learned in step 1. Value of the considered index (starting from lower levels in lattice representation) is adjusted considering the values of its upper and lower neighbors. This is done by computing the following quantities for the considered node $\mu(i)$:

1/mean value of upper neighbors denoted by $m_{up}(i)$

2/mean value of lower neighbors denoted by $m_{low}(i)$

3/minimum distance between considered index and its upper (resp. lower) neighbors, denoted $d_{up}(i)$, (resp. $d_{low}(i)$).

If $m_{up}(i)+m_{low}(i)-2\mu(i)>0$, then $\mu(i)$ is increased:

$$\mu^{new}(i)=\mu^{old}(i)+\beta\frac{(m_{up}(i)+m_{low}(i)-2\mu(i))d_{up}(i)}{2(m_{up}(i)+m_{low}(i))} \quad (2)$$

otherwise $\mu(i)$ is decreased:

$$\mu^{new}(i)=\mu^{old}(i)+\beta\frac{(m_{up}(i)+m_{low}(i)-2\mu(i))d_{low}(i)}{2(m_{up}(i)+m_{low}(i))} \quad (3)$$

β is a constant value in [0, 1]. Several iterations can be done.

For the described patient and the preferences and using the 2 described approaches, the resulting evaluations of the phases of illness are given in Table 7. The results are obtained using the Choquet integral as described in part III, and for HLMS reasoning approach the required index values are obtained after 300 iterations of the described index updating procedure using Java code.

TABLE 7: PHASE EVALUATION TABLE FOR THE PATIENT.

Phase	Propagated interaction approach	HLMS reasoning approach (300 iterations, β=1)
N	0.08	0.08
I	0.4	0.4
II	0.64	0.6251
III	0.64	0.6285

Patients phase is determined based on maximum value of the Choquet integral:

$$Patient\_phase=\max\{Ch_{\mu N},Ch_{\mu I},Ch_{\mu II},Ch_{\mu III}\} \quad (4)$$

In Table 7, phases II and III are determined as equal by the first approach, and the second approach gives a slightly higher score to phase III. As expected, evaluations on phases are similar for the represented example, where the propagated synergy approach was possible, but the second approach is far superior in terms of demands on expert evaluating the patient and in terms of applicability for any given situation. For the phases N and I, the results are identical as expected, since for the described patient, the Choquet integral for those phases does not entail any of the coefficients affected by synergy of two symptoms - PaO_2 and $PaCO_2$.

VIII. CONCLUSION

Fuzzy sets theory has been well studied. Even so, fuzzy measure identification is still a challenging problem. We have shown one approach to solving our problem and shown the results. Further improvements in fuzzy measures identification could come from similar algorithms, taking into account the unique structure of fuzzy measure coefficients and incorporating it into logic of the algorithm. For the presented ARDS classification problem, if machine learning was to take place without the expert evaluation, but based on training data alone, HLMS algorithm could be applied, but even better results could be expected from a more complex fuzzy measure identification than one accomplished by the HLMS algorithm. This is due to the fact that the medical expert determines the interaction of symptoms based on their values for the given patient, and standard online learning of indexes based on training set would not suffice, but the relation of the values of symptoms would also have to be taken into account.

LITERATURE

[1] D. Dubois and H. Prade. A review of fuzzy set aggregation connectives. *Information Sciences*, 36:85.121, 1985.

[2] J.C. Fodor and M. Roubens. *Fuzzy Preference Modelling and Multi-Criteria Decision Aid*. Kluwer Academic

[3] R.R. Yager. Connectives and quantifiers in fuzzy sets. *Fuzzy Sets & Systems*, 40:39.75, 1991.

[4] M. Sugeno: *Theory of fuzzy integrals and its applications*. PhD thesis, Tokyo Institute of Technology, 1974.

[5] M. Grabisch: Fuzzy integral in multicriteria decision-making, *Fuzzy Sets and Systems,* 69, 1995., pp. 279-298.

[6] D.Z.Šaletić, M.Savić, Software Implementation of Fuzzy Decision-Making in Determining the Severity of Respiratory Distress, Ed. Aniko Szakal *Proceedings of the 6th Symposium on Intelligent Systems and Informatics,* SISY 2008, September 26-27, 2008, Subotica, Serbia, 27-sisy2008-submission.pdf.

[7] M.Grabisch. A new algorithm for identifying fuzzy measures and its application to pattern recognition. In: *Proc. Joint Conf. 4th IEEE Internat. Conf. on Fuzzy Systems and 2nd Internat. Fuzzy Engineering Symposium*, pp. 145-150.

[8] T. Mori, T. Murofushi, An analysis of evaluation model using fuzzy measure and the Choquet integral, *5th Fuzzy Systems Symposium*, Kobe, Japan, june 2-3 1989.

[9] M. Grabisch, M. Sugeno, Multi-attribute classification using fuzzy integral, *1st FUZZ'IEEE Congress*, San Diego, march 92, 47-54.

[10] D.Z. Šaletić, *Rasplinuti skupovi i rasplinuto-logički sistemi*, Eduka, Beograd, 2012.

[11] M. Savić: *Rasplinuto višekriterijumsko odlučivanje i primena u medicini – u određivanju stepena težine respiratornog poremećaja pacijenta*, Diplomski rad, Računarski fakultet, Beograd, 2008.

ABSTRACT

Rad se bavi problemom identifikacije rasplinutih mera u višekriterijumskom odlučivanju sa kriterijumima u interakciji. Predstavljen je pristup problemu, zasnovan na algoritmu za identifikovanje rasplinutih mera autora M. Grabiša. Opisana je procedura identifikovanja rasplinutih mera zasnovana na potrebama prezentovanog zadatka i prikazan je primer.

IDENTIFIKACIJA RASPLINUTIH MERA ZA VIŠEKRITERIJUMSKO ODLUČIVANJE

Milica Savić