

Sorting multiattribute alternatives: The TOMASO method

Jean-Luc Marichal ^{a,1} Patrick Meyer ^{b,2} Marc Roubens ^b

^a*Département d'Informatique de Gestion, University of Luxembourg, 162A,
Avenue de la Faïencerie, L-1511 Luxembourg, Luxembourg*

^b*Institute of Mathematics, University of Liège, Grande Traverse 12 - B37, Sart
Tilman, B-4000 Liège, Belgium*

Abstract

We analyze a recently proposed ordinal sorting procedure (TOMASO) for the assignment of alternatives to graded classes and we present a freeware constructed from this procedure. We illustrate it by two examples, and do some testing in order to show its usefulness.

Key words: multi-attribute decision-making, ordinal data, interacting points of view, Choquet integral, freeware TOMASO.

1 Introduction

In this paper we analyze an ordinal sorting procedure for the assignment of alternatives to graded classes in the presence of interacting points of view. This procedure has been recently proposed by Roubens [7,9] and is now known under the name TOMASO, which means “Tool for Ordinal Multi-Attribute Sorting and Ordering.”

The technique used in this method proceeds in two steps: a pre-scoring phase determines for each point of view and for each alternative a net score (the number of times a given alternative beats all the other alternatives minus the number of times this alternative is beaten by the others) and is followed by an aggregation phase, using the discrete Choquet integral, which produces a

¹ Corresponding author. Fax: +352-46-66-44-513. Email: marichal[at]cu.lu

² Supported by the project CELoFA 9813803 of the Région Wallonne (Belgium)

global net score associated to each alternative. These global scores are then used to assign the alternatives to graded classes.

The fuzzy measure associated to the Choquet integral can be learnt from a subset of alternatives (called prototypes) that are assigned beforehand to the classes by the decision maker. This leads to solving a linear constraint satisfaction problem whose unknown variables are the coefficients of the fuzzy measure.

Once a fuzzy measure is found, it is useful to interpret it through some behavioral parameters:

- (1) The importance indices (Shapley importance indices), which make it possible to appraise the overall importance of each point of view,
- (2) The interaction indices (Shapley interaction indices), which measure the extent to which the points of view interact (positively or negatively).

The main purpose of this paper is to present a freeware, also called TOMASO, which allows the user to apply this method to a learning set and analyze the resulting assignments. It is written in Visual Basic, which increases its ergonomomy.

The paper is organized as follows. In Section 2 we recall the general sorting procedure TOMASO. In Sections 3–5 we present the freeware TOMASO, its implementation, and its application on two particular case studies. Finally, in Section 6, we compare this procedure with classical methods in terms of both results and computational efficiency.

2 The TOMASO method

In the present section, we briefly recall the procedure TOMASO. For theoretical developments on this procedure, the readers are referred to [7].

Let A be a set of q potential alternatives, which are to be assigned to disjoint classes, and let $N = \{1, \dots, n\}$ be a label set of points of view to satisfy. For each point of view $i \in N$, the alternatives are evaluated according to a s_i -point ordinal performance scale; that is, a totally ordered set

$$X_i := \{g_1^i \prec_i g_2^i \prec_i \dots \prec_i g_{s_i}^i\}.$$

We assume that each alternative $x \in A$ can be identified with its corresponding profile

$$(x_1, \dots, x_n) \in \times_{i=1}^n X_i =: X,$$

where, for any $i \in N$, x_i represents the partial evaluation of x related to point of view i .

Now, consider a partition of X into m nonempty classes $\{Cl_t\}_{t=1}^m$, which are increasingly ordered; that is, for any $r, s \in \{1, \dots, m\}$, with $r > s$, the elements of Cl_r have a better evaluation than the elements of Cl_s .

We also set

$$Cl_r^{\geq} := \bigcup_{t \geq r} Cl_t \quad (r = 1, \dots, m).$$

The procedure TOMASO consists in partitioning the elements of A into the classes $\{Cl_t\}_{t=1}^m$. It is mainly based on the following result, adapted from [4, Theorem 2.1], which states that, under a simple condition of monotonicity, it is possible to find a discriminant function that strictly separates the classes Cl_1, \dots, Cl_m by ordered numerical thresholds.

For any $x_i \in X_i$ and any $y_{-i} \in X_{-i} := \times_{j \in N \setminus \{i\}} X_j$, we set

$$x_i y_{-i} := (y_1, \dots, y_{i-1}, x_i, y_{i+1}, \dots, y_n) \in X.$$

Theorem 1 *The following two assertions are equivalent:*

(1) *For all $i \in N$, $t \in \{1, \dots, m\}$, $x_i, x'_i \in X_i$, $y_{-i} \in X_{-i}$, we have*

$$x'_i \succ_i x_i \text{ and } x_i y_{-i} \in Cl_t \Rightarrow x'_i y_{-i} \in Cl_t^{\geq}.$$

(2) *There exist*

- *functions $g_i : X_i \rightarrow \mathbb{R}$ ($i \in N$), strictly increasing, called criteria,*
- *a function $f : \mathbb{R}^n \rightarrow \mathbb{R}$, increasing in each argument, called discriminant function,*
- *$m - 1$ ordered thresholds $\{z_t\}_{t=2}^m$ satisfying*

$$z_2 \leq z_3 \leq \dots \leq z_m$$

such that, for any $x \in X$ and any $t \in \{2, \dots, m\}$, we have

$$f[g_1(x_1), g_2(x_2), \dots, g_n(x_n)] \geq z_t \Leftrightarrow x \in Cl_t^{\geq}.$$

For a practical use of this result and in order to produce a meaningful result, Roubens [9] restricted the family of possible discriminant functions to the class of n -variable Choquet integrals and the criteria functions to normalized scores.

Such scores, whose definition might vary from an application to another, should have a precise meaning for the decision maker.

Two natural approaches can be considered: either the score of each alternative is built on the basis of all the alternatives in A or this score is constructed in a context-free manner, that is, independently of the other alternatives. The decision maker must be aware that the final results may significantly differ according to the considered approach. Therefore, a prior analysis of the problem is recommended to choose the scores appropriately.

In the first approach, one possible way to build the scores is to consider comparisons of the alternatives on each of the points of view. We consider $S_i(x)$, the i th partial net score of alternative $x \in A$ along point of view $i \in N$, as the number of times that x is preferred to any other alternative of A minus the number of times that any other alternative of A is preferred to x for point of view i . We furthermore normalize these scores so that they range in the unit interval, i.e.,

$$S_i^N(x) := \frac{S_i(x) + (q - 1)}{2(q - 1)} \in [0, 1] \quad (i \in N),$$

where $q = |A|$. Clearly, this normalized score is not a utility, and should not be considered as such. Indeed, observing an extreme value (close to 0 or 1) means that x is rather “atypical” compared to the other alternatives along point of view i . Thus, the resulting evaluations strongly depend on the alternatives which have been chosen to build A .

Consider now the second approach, that is, where the score of each alternative does not depend on the other alternatives in A . In this case, we suggest the decision maker provides the score functions as utility functions. Alternatively, we can approximate these utility functions by the following linear formula:

$$S_i^N(x) := \frac{\text{ord}_i(x) - 1}{s_i - 1} \in [0, 1] \quad (i \in N),$$

where $\text{ord}_i : A \rightarrow \{1, \dots, s_i\}$ is a mapping defined by $\text{ord}_i(x) = r$ if and only if $x_i = g_r^i$. In this latter case, S_i^N does not necessarily represent a real utility and probably does not correspond to the utility the decision maker has in mind. We therefore continue to call it a score.

Notice that the case studies we present in this paper are treated by means of the scores of the first type, i.e., based on the comparison of alternatives.

The normalized scores of each alternative x are then aggregated by means of a Choquet integral [2], namely

$$\mathcal{C}_v(S^N(x)) := \sum_{i=1}^n S_{(i)}^N(x) [v(A_{(i)}) - v(A_{(i+1)})],$$

where $S^N(x)$ stands for $(S_1^N(x), \dots, S_n^N(x))$ and v represents a fuzzy measure

on N ; that is, a monotone set function $v : 2^N \rightarrow [0, 1]$ fulfilling $v(\emptyset) = 0$ and $v(N) = 1$. This fuzzy measure merely expresses the importance of each subset of points of view. Also, the parentheses used for indices represent a permutation on N such that

$$S_{(1)}^N(x) \leq \dots \leq S_{(n)}^N(x),$$

and, for any $i \in N$, $A_{(i)}$ represents the subset $\{(i), \dots, (n)\}$.

The Choquet integral presents standard properties for aggregation (see [3,6]): it is continuous, non decreasing, located between min and max. The major advantage linked to the use of this integral derives from the large number of parameters ($2^n - 2$) associated with a fuzzy measure. However, this flexibility can be also considered as a serious drawback when assessing the fuzzy measure. To reduce the number of parameters, it is proposed to consider k -additive fuzzy measures [3]: a fuzzy measure v on N is k -additive if its Möbius transform $m : 2^N \rightarrow \mathbb{R}$, defined by

$$m(S) = \sum_{T \subseteq S} (-1)^{|S|-|T|} v(T) \quad (S \subseteq N),$$

satisfies $m(S) = 0$ for S such that $|S| > k$ and there exists at least one subset S such that $|S| = k$ and $m(S) \neq 0$. It can be shown that k -additive fuzzy measures on N can be represented by at most $\sum_{i=1}^k \binom{n}{i}$ parameters.

We now explain how the fuzzy measure is assessed in this procedure.

Assume that all the alternatives of $A \subseteq X$ are already sorted into classes Cl_1, \dots, Cl_m . In some particular cases there exist a fuzzy measure v on N and $m - 1$ ordered thresholds $\{z_t\}_{t=2}^m$ satisfying

$$z_2 \leq z_3 \leq \dots \leq z_m$$

such that for any $x \in A$, and any $t \in \{2, \dots, m\}$, we have

$$\mathcal{C}_v(S^N(x)) \geq z_t \Leftrightarrow x \in Cl_t^{\geq}.$$

Of course, if such a fuzzy measure does exist then the thresholds may be defined by

$$z_t := \min_{x \in Cl_t^{\geq}} \mathcal{C}_v(S^N(x)) \quad (t = 2, \dots, m).$$

In real situations, the assignment of all alternatives is not known but has to be determined. However, this assignment, or equivalently the fuzzy measure v , can be learnt from a reference subset, made up of prototypes that have been sorted beforehand by the decision maker.

Practically, the decision maker is asked to supply such a prototypical subset $P \subseteq A$ and the assignment of each of its elements to a given class; that is, a partition of P into prototypical classes $\{P_t\}_{t=1}^m$, where $P_t := P \cap Cl_t$ for all $t \in \{1, \dots, m\}$. Here some prototypical classes may be empty.

As the Choquet integral is supposed to strictly separate the classes Cl_t , we must impose the following necessary condition

$$\mathcal{C}_v(S^N(x)) - \mathcal{C}_v(S^N(x')) \geq \varepsilon, \quad (1)$$

for each ordered pair $(x, x') \in P_t \times P_{t-1}$ and each $t \in \{2, \dots, m\}$, where ε is a given strictly positive threshold.

These separation conditions, put together with the boundary and monotonicity constraints on the fuzzy measure, form a linear constraint satisfaction problem whose unknowns are the coefficients of the fuzzy measure. Thus the sorting problem consists in finding a feasible solution satisfying all these constraints. If ε has been chosen too big, the problem might have no solution. To avoid this, we can consider ε as a non-negative variable to be maximized. In this case its optimal value must be strictly positive for the problem to have a solution.

In the resolution of this problem, we use the principle of parsimony. If no solution is found for $k = 1$, we turn to $k = 2$. If no solution is still found, we turn to $k = 3$, and so forth, up to $k = n$. Notice however that an empty solution set for $k = n$ is necessarily due to an incompatibility between the assignment of the given prototypes and the assumption that the discriminant function is a Choquet integral.

Due to the increasing monotonicity of the Choquet integral, the number of separation constraints (1) can be reduced significantly. For example, if $x'' \in P_{t-1}$ is such that $\mathcal{C}_v(S^N(x')) \geq \mathcal{C}_v(S^N(x''))$ then, by transitivity, the constraint

$$\mathcal{C}_v(S^N(x)) - \mathcal{C}_v(S^N(x'')) \geq \varepsilon$$

is redundant.

Now, we can define a dominance relation D on X as follows: For each $x, y \in X$,

$$xDy \Leftrightarrow S_i^N(x) \geq S_i^N(y) \quad \forall i \in N.$$

Being an intersection of complete (total) orders, the binary relation D is a partial order, i.e., it is reflexive, antisymmetric, and transitive. Furthermore we clearly have

$$xDy \Rightarrow \mathcal{C}_v(S^N(x)) \geq \mathcal{C}_v(S^N(y)).$$

It is then useful to define, for each $t \in \{1, \dots, m\}$, the set of non-dominating alternatives of P_t ,

$$Nd_t := \{x \in P_t \mid \nexists x' \in P_t \setminus \{x\} : xDx'\},$$

and the set of non-dominated alternatives of P_t ,

$$ND_t := \{x \in P_t \mid \nexists x' \in P_t \setminus \{x\} : x'Dx\},$$

and to consider only constraint (1) for each ordered pair $(x, x') \in Nd_t \times ND_{t-1}$ and each $t \in \{2, \dots, m\}$. Thus, the total number of separation constraints boils down to

$$\sum_{t=2}^m |Nd_t| |ND_{t-1}|.$$

Finally, suppose that there exists a k -additive fuzzy measure v^* that solves the problem above. Then any alternative $x \in A$ will be assigned to

- the class Cl_t if

$$\min_{y \in Nd_t} \mathcal{C}_{v^*}(S^N(y)) \leq \mathcal{C}_{v^*}(S^N(x)) \leq \max_{y \in ND_t} \mathcal{C}_{v^*}(S^N(y)),$$

- one of the classes Cl_t or Cl_{t-1} if

$$\max_{y \in ND_{t-1}} \mathcal{C}_{v^*}(S^N(y)) < \mathcal{C}_{v^*}(S^N(x)) < \min_{y \in Nd_t} \mathcal{C}_{v^*}(S^N(y)).$$

3 The freeware TOMASO

We now present the freeware TOMASO,³ which is an implementation (written in Visual Basic) of the algorithm presented in Section 2. In the present section we describe briefly the various aspects of the freeware. In the next section, we present an application on a fictitious example of evaluation of students, we show how the freeware deals with a leave-one-out procedure, and we present a way to test its performance on this particular example. In the following section, we present an application of the method to a classification problem of quality indices for association rules.

The freeware TOMASO can be used for supervised classification of ordinal multi-criteria data. In order to work properly, it requires some information about the structure of the data. We call a prototypical set the set of objects used to build the classifier. It is described by a set of alternatives which are already classified by the decision maker.

³ Available at <http://patrickmeyer.tripod.com>

This means that the class of problems which can be solved by this method is quite particular. The decision maker must be aware of some a priori information on certain objects of his decision problem. He must be an expert in the field and should be able to give a global evaluation on some particular elements. Later, when the model is built, new alternatives can be considered. The classifier will then assign them to one or more of the predefined classes, according to the preferences of the decision maker.

When using the freeware, at first the user has to load the file with the data he wants to analyze⁴. These data have to be of ordinal nature and, as mentioned before, can be composed of the prototypes and some alternatives that need to be classified.

After the calculation of the net scores, the user has to fix the number of classes and their structures. This last point is achieved by adding the prototypes to their respective classes, which can be done either manually or by loading a file. At the end of this step, the alternatives which are not assigned to any class do not belong to the prototypical set and will either be used as a test subset (to check the validity of the method) or simply as objects to be classified.

The next stage of the freeware is to check the consistency of the assignments within the set of prototypes. The inconsistencies can be of two types:

- Two alternatives x and y with the same net scores for each criterion ($S^N(x) = S^N(y)$) belong to two different classes
- Two alternatives x and y so that x dominates y (i.e. $S_i^N(x) \geq S_i^N(y) \forall i \in N$) and y belongs to a better class than x .

If no inconsistency is detected, the freeware allows the user to go on. Else, the user has to change the definition of the classes by the prototypes (either manually, or by loading another class file). The next step is the determination of the non-dominating set of alternatives and the non-dominated set of alternatives in each class. These sets are not empty, because the user should add at least one alternative in each class. As already mentioned earlier, this is meant to reduce the number of separation constraints and is justified by the increasing monotonicity of the Choquet integral.

The resolution of the linear constraint satisfaction problem is the next stage. The user can choose between fixing the value of $k < n$ or solving the problem for $k = n$. In both cases, a file of constraints is computed. An external solver⁵ uses this file as an input and tries to solve the problem.

⁴ See tutorial at <http://patrickmeyer.tripod.com> for a detailed description of the data file formats

⁵ `lp_solve` 3.0, downloadable at ftp://ftp.ics.ele.tue.nl/pub/lp_solve/, released under the LGPL license

If the problem has no solution for $k = n$, the user should revise the definition of the classes. On the other hand, if there is a solution for the problem, the following results are displayed:

- The coefficient of the fuzzy measure v ,
- The values of the Möbius transform m of v ,
- The value of ε ,
- The borders of the classes (for each class Cl_t , given by the maximal and the minimal value of the Choquet integral of the alternatives of Cl_t),
- The Shapley importance indices for each point of view i ,
- The Shapley interaction indices for each pair of points of view,
- The Choquet integral of each prototype.

An important tool is the assignment of the alternatives which are not in the prototypical set to the classes, according to the previously built model. During this step, it may happen that the Choquet integral of some of these objects may not be between the limits of any of the classes. In that case, those alternatives are assigned to “imprecise” classes which lie between the original classes. If for one particular alternative x the decision maker is not satisfied with such a fuzzy assignment between classes Cl_r and Cl_s , he may revise the definition of the adjacent classes in order to include x or maybe another alternative similar to x . Else, one can say that alternative x simply belongs to class Cl_r or Cl_s .

To check the structure of the set of prototypes, it is possible to apply a particular leave-one-out procedure to the data. For each alternative x of the prototypical set, the model is rebuilt without the alternative. Then x is assigned to a class, according to the prototypes. This class should then be the original class of x . If not, it is considered as an error. At the end of the whole procedure, a high error ratio, say e/p , with e close to p (p is the cardinality of the prototypical set and e is the number of badly classified alternatives during the leave-one-out procedure) does not necessarily mean that the data of the prototypes are badly chosen. It either stands for a “minimal” prototypical set, where nearly each alternative is important for the building of the model, or for a quite complex data structure. In this latter case, the set of prototypes should perhaps be revised and enriched with new alternatives to increase its diversity.

4 Assigning students to graded classes

In order to validate the TOMASO method, we apply it on the following example. Consider a set of students evaluated in three courses: mathematics, physics, and literature. For each course, the evaluation scale has three ordered qualitative levels: bad \prec medium \prec good. In total this makes 27 possible dif-

ferent students. Besides, for each student, a decision maker has given a global evaluation on an qualitative ordinal scale with three levels: bad (1) \prec medium (2) \prec good (3). We analyze two different subproblems of this example. The first one uses the complete set of 27 students as prototypes and any future assignment has to be correct, because the 27 students represent all possible cases. The second study is on a random subset of the set of 27 students. It is somewhat more realistic, as a global information is known only on a subset of students. There, future assignments are not necessarily correct. In order to test the usefulness of the method, a particular algorithm of cross validation will be applied to the problem.

4.1 Subproblem 1: 27 students

Figure 1 represents the assignments of the decision maker, and the dominance relations between students of the same class. The determination of the non-

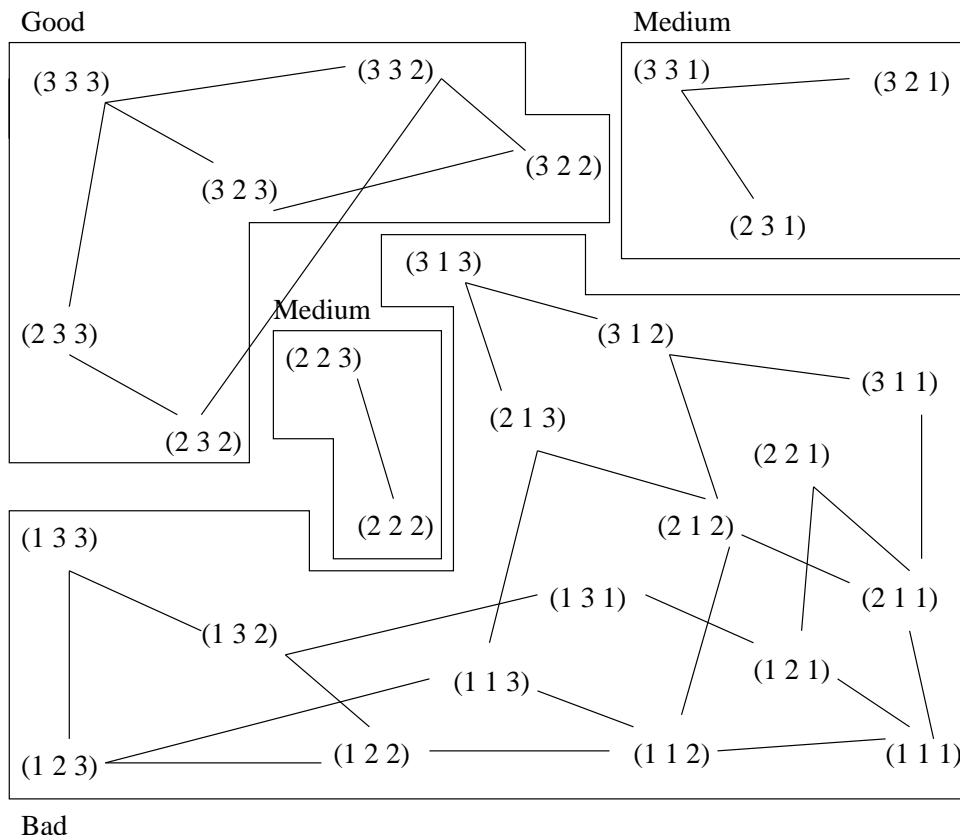


Fig. 1. Assignments of the decision maker

dominating and non-dominated sets of alternatives for each class results in the following table:

ND_g	(333)		
ND_m	(223)	(331)	
ND_b	(133)	(313)	(221)
Nd_g	(232)	(322)	
Nd_m	(222)	(231)	(321)
Nd_b	(111)		

The extreme sets (ND_g and Nd_b) will not be used in the linear constraint satisfaction program.

The problem has a solution for $k = n = 3$, with $\varepsilon = 0.086$ (but no solution can be found for $k < 3$). This means that all the prototypes can be assigned to their correct classes with the help of a Choquet integral as a discriminant function. Interesting parameters are the importance indices which are shown in the following table:

Math	Physics	Literature
0.417	0.417	0.166

4.2 Subproblem 2: a subset of the 27 students

For this problem, we chose a random subset of the 27 prototypical students as follows: we scan sequentially the original set of 27 students and assign each student with a probability of 1/2 to the new subset. The randomly generated subset used here is described in the following table:

Good	(322)	(332)	(323)	(333)				
Medium	(321)	(331)						
Bad	(311)	(121)	(131)	(312)	(132)	(313)	(123)	(133)

The structure of this problem is given in Figure 2. The original classes are represented by dashed lines. The prototypical classes are described by solid lines. The goal is to reconstruct the original classes as accurately as possible.

The determination of the non-dominating and non-dominated sets of alternatives for each class results in the following table:

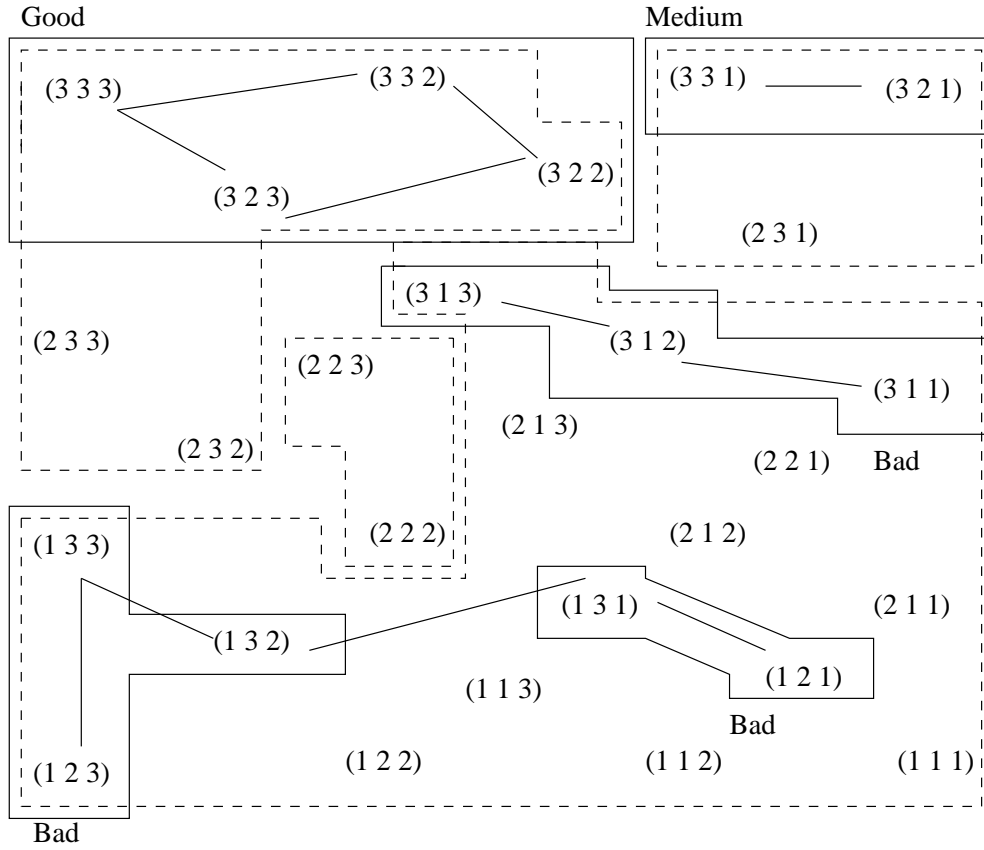


Fig. 2. Assignments of the decision maker

ND_g	(333)
ND_m	(331)
ND_b	(133) (313)
Nd_g	(322)
Nd_m	(321)
Nd_b	(311) (121) (123)

We observe that these sets are quite reduced in comparison to the previous subproblem.

The problem has a solution for $k = n = 3$, with $\varepsilon = 0.115$. The importance indices are given in the following table:

Math	Physics	Literature
0.389	0.389	0.222

They are similar to those of the first subproblem. Mathematics and Physics have the highest importance indices with equal values.

We can also see that there is a solution for $k = 2$, with $\varepsilon = 0.069$. In this case, the importance indices are given in the following table:

Math	Physics	Literature
0.500	0.300	0.200

The (pairwise) interaction indices are given in the following table:

(Math, Physics)	(Math, Literature)	(Physics, Literature)
0.200	0	0.400

Coming back to the solution with $k = 3$, the leave-one-out procedures gives a misclassification ratio of 14.29%. This means that 2 alternatives out of the 14 of the prototypes were badly reclassified. Besides, the importance indices stay rather constant during these 14 leave-one-out procedures. The three next tables resume the value of the Shapley indices for the three branches (mathematics (m), physics (p) and literature (l)).

m	0.389	0.389	0.375	0.389
p	0.389	0.389	0.375	0.389
l	0.222	0.222	0.250	0.222

m	0.417	0.389	0.389	0.389	0.389
p	0.417	0.389	0.389	0.389	0.389
l	0.167	0.222	0.222	0.222	0.222

m	0.389	0.389	0.389	0.389	0.389
p	0.389	0.389	0.389	0.389	0.389
l	0.222	0.222	0.222	0.222	0.222

The average values for mathematics and physics are both 0.39 with a variance of $7.43 \cdot 10^{-5}$. For literature the mean is 0.220 with a variance of $2.89 \cdot 10^{-4}$.

The assignments of the remaining 13 alternatives are resumed in the following table:

Student	Assignment	Original Class
(111)	bad	bad
(211)	bad	bad
(221)	medium	bad
(231)	medium	medium
(112)	bad	bad
(212)	bad	bad
(122)	bad	bad
(222)	good	medium
(232)	good	good
(113)	bad	bad
(213)	bad	bad
(223)	good	medium
(233)	good	good

The three erroneous assignments (indicated in boldface type) lead to an error rate of 23.08%. If we analyze these problems, they are not surprising at all. The randomly chosen subset has very few representatives in the middle class of medium students. The three erroneous assignments slide from the middle to the top class, or from the lower to the middle class. Hence we can conclude that the class of medium students is quite problematic in this particular case. But the error rate is still satisfactory.

In order to determine if the algorithm used in the TOMASO method is useful on this particular example of 27 students, we use a special 2-fold cross validation. The goal is to prove that the previous randomly chosen subproblem is not a particular case, and that pertinent conclusions can be drawn for any possible prototypical subset of the 27 students.

The general idea behind this test is to split the data D of the 27 students in 2 random subsets, a training set R and a test set T . The model is built on R and tested on T . The resulting classification (on T) is then compared to the original classes of D . To determine the accuracy of the model, an error ratio is calculated by dividing the number of misclassified elements of T by the cardinality of T .

In this particular example, we build a certain number S of subsets of the set of 27 students. One condition needs to be fulfilled: there must be at least one

object of each decision class in the subset. To do so, we scan sequentially the original set of 27 students and assign each student with a probability of 50% to the new subset. Afterwards, we check if each class is represented at least once in the new set of students.

Let X be a random variable which represents the error ratio. Our main concern is the mean value of X , $E[X]$, on all possible subsets of D under the constraint described before.

In order to give an estimation of $E[X]$ we calculate the mean \bar{X} of X on 30 subsets. The values for the X_i are given in the following table:

0.2857	0	0.2308	0.2500	0.5455
0.1538	0	0.1429	0.0667	0.1818
0.3571	0.4286	0	0.4286	0.5333
0.3889	0.3333	0	0.1538	0.0769
0.2500	0.2941	0.2308	0.3077	0.0769
0.1250	0.1333	0.2222	0.3125	0.2308

Hence $\bar{X} = 0.2247$ and the standard deviation $s = 0.0231$.

Besides, the normality of the distribution given by the values in the previous table cannot be rejected. Hence we compute the Student confidence interval around the value of \bar{X} . Finally:

$$P(E[X] \in [0.2131, 0.2363]) = 99\%.$$

More intuitively, we can claim that with a probability of 99%, TOMASO allows to classify the students of this example with an error rate between 21.31% and 23.63% if the prototypes are chosen randomly as described beforehand.

We observe that the previous example treated in Subsection 4.2 is not aberrant, and that it could represent a real case study. Further more the conclusions drawn from it are pertinent.

We also have analyzed the variation of the importance indices over these 30 cross validations. In a majority of cases, the order “mathematics \sim physics \succ literature” is maintained. But the values for each criterion are varying significantly. It is important to emphasize here that the exact values of these indices should merely be considered as parameters.

5 Assigning quality measures of association rules to graded classes

We present a second example to show that the TOMASO method can be used in various real-life problems where a MCDA problem requires graded classes as an output. This particular example comes from the data mining and knowledge extraction fields. Certain data mining algorithms, especially those used for unsupervised learning, can generate a large number of rules ($A \rightarrow B$; read: if A is satisfied then B is probably also satisfied). They need to be evaluated before being presented to an expert in order to reduce their number, and to retain only the “best” and the most interesting ones. Intuitively one could say that an association rule is “good” if it has few counterexamples. Therefore, many quality measures have been proposed to evaluate rules. Most of them have different properties and express more subtle properties than simply counting the counterexamples. These properties are often contradictory and conflicting. In [5] the authors have described quality measures on a limited set of criteria. Currently, the work has progressed and a set of 20 measures is evaluated on 8 criteria. The interested reader should refer to [5] for further details. The list of criteria is given hereafter:

- g_1 : Asymmetric processing of A and B (To distinguish $A \rightarrow B$ rules from $B \rightarrow A$ rules) (asymmetric (1) \succ symmetric (0)).
- g_2 : Decrease with n_b ($n_b = |B|$) with fixed $|A|$ (decreasing with n_b (1) \succ non-decreasing with n_b (0)).
- g_3 : Constant at independence (If A and B are independent, the rule contains no information. Hence the quality measure should take a fixed value at independence (e.g. 0)) (fixed value at independence (1) \succ value at independence varying with rule (0)).
- g_4 : Constant at situation of logical rule (If no counterexample to the rule exists, the measure should take a fixed maximum value (e.g. 1)) (constant at maximum (1) \succ maximum value dependant on rule (0)).
- g_5 : Shape of the curve [measure = $f(\text{counterexamples})$] (A slow decrease with the number of counterexamples allows to keep the rule longer, a fast decrease with the number of counterexamples bans the rule faster; depends on the preferences of the expert) (concave (2) \succ linear (1) \succ (0) convex or convex (0) \succ linear (1) \succ concave (2)).
- g_6 : Sensitivity to the total number of examples of the analyzed data table (sensible (1) \succ not sensible (0)).
- g_7 : Ease to fix an acceptance threshold (If the value of the measure lies above this threshold, the rule is accepted, else it is rejected) ((1) easy \succ hard (0)).
- g_8 : Intelligibility of the measure (easily explainable (2) \succ basic mathematical knowledge required (1) \succ hardly explainable (0)).

The data matrix is shown in Table 1. The expert would like to assign these 20

	g_1	g_2	g_3	g_4	g_5	g_6	g_7	g_8
SUP	0	0	0	0	1	0	1	2
CONF	1	0	0	1	1	0	1	2
R	0	1	1	0	1	0	1	1
CENCONF	1	1	1	0	1	0	1	2
PS	0	1	1	0	1	1	1	1
LOE	1	1	1	1	1	0	1	1
ZHANG	1	1	1	1	2	0	0	0
- IMPIND	1	1	1	0	1	1	1	0
LIFT	0	1	1	0	1	0	1	1
SURP	1	1	0	0	1	0	1	1
SEB	1	0	0	1	0	0	1	1
OM	1	1	1	1	0	0	1	2
CONV	1	1	1	1	0	0	1	1
ECR	1	0	0	1	2	0	1	1
KAPPA	0	1	1	0	1	0	1	0
IG	0	1	1	0	2	0	1	0
INTIMP	1	1	1	1	2	1	1	0
EII	1	1	1	1	2	1	0	0
PDI	1	1	1	0	1	1	1	0
LAP	1	0	0	0	1	0	1	0

Table 1
Decision matrix for the quality measures problem

measures to 3 graded classes: Good \succ Medium \succ Bad. During a discussion, he is asked to provide at least one measure for each class. He chooses a few measures he knows quite well, and for which he is able to interpret the resulting rules and their quality. The assignment of the prototypes is given in Table 2. One should mention here that this assignment is reflecting the preferences of

Good	CENCONF	LOE	OM	EII
Medium	ZHANG	ECR	PDI	
Bad	SUP	SURP	KAPPA	

Table 2
Assignment of the prototypes

the expert. He is not necessarily a data mining expert. Therefore, he is not

aware of all the subtleties of the measures, but during the discussion, he gets information on the 8 selected properties and on certain measures.

A solution is found for $k = 1$. This additive model allows to classify the non-prototypical alternatives as shown in Table 3. The importance indices are

Good	CONF	CONV	INTIMP
Medium	PS	-IMPIND	SEB
Medium \cup Bad	R	LIFT	
Bad	LAP	IG	

Table 3

Assignment of the remaining measures

shown in Table 4. These results are in close relation with the current research which is made on the evaluation of quality measures of decision rules. In [5] the authors analyze the same type of problem, with the Promethee-Gaia method for multiple criteria ranking. Our results are not in contradiction with those published in [5].

6 Benefits and specificities of TOMASO and comparison to two classical MCDA methods

In the general context of multicriteria decision aiding and sorting methods, TOMASO has two major particular features.

First of all, it is able to cope with interacting criteria. This allows to deal with a larger set of problems, and the definition of the criteria is less restrictive than for classical methods (like PROMETHEE [1], or ELECTRE TRI [8]).

Secondly, the decision maker doesn't have to provide difficult information on parameters of the model (like weights for the criteria). They are "learnt" from a set of well-known alternatives by the decision maker. One could say that in certain problems this may be as complicated as determining the weights of the criteria. But our experience has shown that in our case studies, expert decision makers have an a priori knowledge about a few objects of the problem they are analyzing. This particular information is fully exploited in TOMASO through the concept of prototypes.

g_1	g_2	g_3	g_4	g_5	g_6	g_7	g_8
0.182	0.000	0.182	0.182	0.000	0.181	0.001	0.273

Table 4

Importance indices

In the theoretical example of the assignment of students to graded classes, we observe first of all that the particular classification of the prototypes implies interaction among criteria (no model could be found for $k = 1$). In case of the two classical MCDA methods PROMETHEE and ELECTRE TRI, the criteria require to be independent. This restrictive condition implies that here a solution which is compatible with the decision maker's preferences cannot be found. The strong condition of independent criteria may be quite restrictive in many real-life cases.

Another point concerns the information that the decision maker has to provide. In PROMETHEE, a weight system has to be determined in order to obtain a final ranking on the alternatives. Some visualization tools allow an easier fixing of these parameters, but nevertheless, this information is quite complex and hard to obtain. In ELECTRE TRI, a procedure allows to provide some prototypical alternatives to the method, in order to fix certain parameters. But other parameters still need human interaction (thresholds). Anyway, nothing guarantees that a solution to the classification of the prototypical alternatives will be found if the criteria are not independent. In TOMASO the decision maker, such as a teacher in a student evaluation, has to express a global feeling on a subset of students. No further parameter has to be fixed. These assignments are "natural" opinions of the teacher on some students, without any precise considerations on the exact weights of the different matters. In this case, this simplified information is enough to build a model, and to provide the "weights" of the criteria by means of the Shapley indices.

A further specificity of the TOMASO method is the possibility to use and work with purely ordinal data. No further information is required. In PROMETHEE, the user has to decide what type of preference function he wants to use in order to make different ordinal and cardinal scales comparable. Besides, each preference function requires some parameters to be fixed by the user. The method ELECTRE TRI also needs some parameters to be fixed by the decision maker. In our case, the problem is solely analyzed (and solved) by asking the decision maker a minimal number of questions on global evaluations concerning the alternatives and no further information on the structure of the data.

7 Conclusion

We have introduced a freeware that uses the procedure TOMASO to build a supervised classifier.

Some interesting observations can be made with the help of this tool. We want to emphasize that the interaction indices as well as the Shapley indices are mainly parameters of the method. In particular the importance indices should

not be interpreted as “real” weights of the problem. Even if they are stable during the leave-one-out procedures, this doesn’t indicate that they “fit” to the decision maker’s mind. As we have shown in the last part of Section 4.2, they can vary quite significantly from one prototypical set to another. At most their order can give an indication on the ranking of the “real” weights.

References

- [1] J.P. Brans and B. Mareschal, *PROMETHEE-GAIA - Une méthode d’aide à la décision en présence de critères multiples* (Ellipse, 2002).
- [2] G. Choquet, Theory of capacities, *Annales de l’Institut Fourier* **5** (1953) 131–295.
- [3] M. Grabisch, k -order additive discrete fuzzy measures and their representation, *Fuzzy Sets and Systems* **92** (1997) 167–189.
- [4] S. Greco, B. Matarazzo, and R. Slowinski, Conjoint measurement and rough set approach for multicriteria sorting problems in presence of ordinal criteria, in: A. Colorni, M. Paruccini, and B. Roy (eds.), *AMCDA: Aide Multicritère à la Décision (Multiple Criteria Decision Aiding)*, European Commission Report, EUR 19808 EN, Joint Research Centre, Ispra, 2001, pp. 117–144.
- [5] Ph. Lenca, P. Meyer, P. Picouet, and B. Vaillant, Aide multicritère à la décision pour évaluer les indices de qualité des connaissances, *RSTI-RIA (EGC 2003)*, 1(17), pp. 271–282, 2003.
- [6] J.-L. Marichal, An axiomatic approach of the discrete Choquet integral as a tool to aggregate interacting criteria, *IEEE Transactions on Fuzzy Systems* **8** (6) (2000) 800–807.
- [7] J.-L. Marichal and M. Roubens, On a sorting procedure in the presence of qualitative interacting points of view, in: J. Chojcan, J. Leski (eds.), *Fuzzy sets and their applications* (Silesian University Press, Gliwice, 2001) pp. 217–230 [ISBN 83-88000-64-0]. Also in: *22nd Linz Seminar on Fuzzy Set Theory (LINZ 2001)*, Linz, Austria, February 6-10, 2001, pp. 64–75.
- [8] V. Mousseau, R. Slowinski, and P. Zielniewicz, ELECTRE TRI 2.0a, Methodological guide and user’s documentation. Document du LAMSADE no.111, Université de Paris-Dauphine, 1999.
<http://www.lamsade.dauphine.fr/mcda/biblio/Author/MOUSSEAU-V.html>
- [9] M. Roubens, Ordinal multiattribute sorting and ordering in the presence of interacting points of view, in: D. Bouyssou, E. Jacquet-Lagrèze, P. Perny, R. Slowinski, D. Vanderpooten, and P. Vincke (eds.), *Aiding Decisions with Multiple Criteria: Essays in Honor of Bernard Roy*, (Kluwer Academic Publishers, Dordrecht, 2001), pp. 229–246.