Aggregation theory revisited

Raúl Pérez-Fernández and Bernard De Baets

***Abstract*—The field of aggregation theory aims at formalizing in a mathematical way the process of combining several inputs into a single output, typically both the inputs and the output being elements of a poset. Although the field in itself only dates from the second half of the last century, one could easily trace back further in time prominent examples of aggregation functions. For instance, means were already studied by Cauchy in the early 1820s. Although the most popular aggregation processes have historically been those dealing with real numbers, the interest of practitioners in different types of data is not to be neglected. Some prominent examples are the aggregation of strings, which is nowadays a popular topic for computer scientists and bioinformaticians, and the aggregation of rankings, which has been studied in social choice theory since the eighteenth century. In this paper, we propose to abandon the current order-based understanding of an aggregation process and embrace a new geometrically-oriented sense upon which a new theory of aggregation could be developed.**

***Index Terms*—Data aggregation; Monotonicity; Betweenness relation; Beset.**

I. Introduction

The notion of a mean or averaging function has for a long time been considered in mathematics and can be traced back to Ancient Greece [1]. The first modern definition is usually attributed to Cauchy in 1821 [2] and just requires the output to be bounded by the minimum and the maximum of the inputs. More recently, the intuitive property of monotonicity, which requires that an increase in the inputs must imply an increase in the output, has been added to the definition that is currently considered standard [3].

Although aggregation was initially considered to be an operation on the real numbers, aggregation on bounded posets has received increasing attention in recent times [4], [5]. Unsurprisingly, the standard definition of an aggregation function on a bounded poset builds again upon the property of monotonicity, while incorporating the property of preservation of the bounds.

As intuitive and general as this definition may sound, the notion of monotonicity has recently withstood great criticism from the scientific community. One only needs to think of the mode in order to find an example of 'aggregation function' that, actually, is not an aggregation function, since it is not monotone. Again in the setting in which the considered bounded poset is a compact real interval, the property of weak monotonicity has been introduced in order to accommodate the mode and other non-monotone functions [6], such as Lehmer means [7] and Gini means [8]. One could note that weak monotonicity generalizes both monotonicity and shift-invariance. A further generalization of the property of weak monotonicity has also been studied under the name of directional monotonicity [9].

In the setting of multidimensional data, monotonicity has also been proved to be problematic. In particular, monotonicity combined with the property of orthogonal equivariance restricts idempotent functions to the family of weighted centroids [10]. This result still holds even when replacing monotonicity and orthogonal equivariance by the weaker property of orthomonotonicity [11]. Another intuitive generalization of the property of monotonicity in the setting of multidimensional data is that of componentwise monotonicity [10]. Unfortunately, this definition also carries an undesirable behaviour since it inevitably leads to componentwise functions in which all components are treated independently. A taxonomy of monotonicity properties for the aggregation of multidimensional data can be found in [12].

The property of monotonicity, understood in an admittedly different (but still related) sense, has also been considered in the field of social choice theory for the aggregation of rankings. To the best of our knowledge, the first formalization of this property can be traced back to the seminal work by Arrow [13] under the name of 'positive association of social and individual values': "If an alternative social state x rises or does not fall in the ordering of each individual without any other change in those orderings and if x was preferred to another alternative y before the change in individual orderings, then x is still preferred to y." However, although always desirable, this monotonicity property has also been proved not to be satisfied by some prominent aggregation methods, as discussed by, for instance, Smith [14] and Fishburn [15].

In other domains addressing the problem of data aggregation, the property of monotonicity has been totally ignored, probably due to the absence of a natural order on most of the considered structures. For instance, the median procedure for the aggregation of binary relations has been a popular topic [16], [17], [18] with no mention of the property of monotonicity. Admittedly, the set of binary relations is a bounded poset where set inclusion acts as the order relation and, interestingly, the aforementioned median procedure turns out to be monotone with respect to this order relation. Nevertheless, when one restricts the attention to specific families of binary relations (which definitely represents the most interesting setting), this poset structure might be lost (e.g., there exists no pair of different rankings included in one another). This probably explains why the study of the property of monotonicity has been historically ignored. The aggregation

Both authors are affiliated with the research unit KERMIT at the Department of Data Analysis and Mathematical Modelling, Ghent University, Coupure links 653, 9000 Gent, Belgium (email: raul.perezfernandez@ugent.be; bernard.debaets@ugent.be). Raúl Pérez-Fernández is also affiliated with the Department of Statistics and O.R. and Mathematics Didactics, University of Oviedo, Oviedo, Spain (email: perezfernandez@uniovi.es). Raúl Pérez-Fernández acknowledges the support of the Research Foundation of Flanders (FWO17/PDO/160) and the Spanish MINECO (TIN2017-87600-P).

of strings, featuring the median string [19] and the closest string [20] problems, is also a quite prominent example of aggregation problem that has been commonly addressed with no mention of the property of monotonicity.

From a more practical point of view, one could think of the aggregation of compositional data. Two liquids of known compositions are mixed (with known mixing ratio) into a new liquid. This new liquid has an associated composition that can be understood as the result of aggregating the compositions of the two original liquids. Unfortunately, this would never be considered an aggregation process since the set of all possible compositions is not naturally ordered (letting aside the Lorenz order [21] and related notions [22] that are obviously not suitable for the purpose of this aggregation process).

In this paper, we follow the direction started in [23] and aim at generalizing the property of monotonicity by moving away from its order-theoretical origin and embracing a new geometrically-oriented meaning arising from the preservation of a betweenness relation. The remainder of the paper is structured as follows. In Section II, we recall some relevant notions concerning the standard aggregation on bounded posets. In Section III, we introduce the notion of an aggregation function on a set equipped with a betweenness relation (beset, for short) that generalizes the standard definition of an aggregation function on a bounded poset. Section IV is devoted to the aggregation of compositional data and the aggregation of binary strings, which are shown to be examples of aggregation functions on a beset. Two examples of commonly-studied aggregation processes that do not fit within this newly-introduced definition of an aggregation function are given in Section V. We end with a discussion on the perspectives arising from this work in Section VI.

II. The standard for aggregation

Means are one of the most common and ancient notions in mathematics [3]. Their early usage was originally restricted to real numbers and just required the output to be bounded by the minimum and the maximum of the inputs. Nowadays, means are additionally required to be monotone and have been generalized in order to deal with any poset, not just the real line [24]. The term averaging function is recently gaining popularity for referring to a mean.

Definition 1: Consider a poset $(X, \leq)$ and $n \in \mathbb{N}$. A function $A : X^n \to X$ on $(X, \leq)$ is called an (n-ary) mean or averaging function on $(X, \leq)$ if

(i) it is idempotent, i.e., $A(x, \ldots, x) = x$, for any $x \in X$;
(ii) it is monotone[1], i.e., for any $\mathbf{x}, \mathbf{y} \in X^n$, the fact that[2] $\mathbf{x} \leq_n \mathbf{y}$ implies that $A(\mathbf{x}) \leq A(\mathbf{y})$.

We recall that, for a monotone function, both idempotence and being bounded from below by the minimum and from above by the maximum are equivalent.

[1]The term 'monotone' actually refers to two types of behaviour: monotone increasing (isotone) and monotone decreasing (antitone). However, for historical reasons, we adhere to the term 'monotone' for referring to 'monotone increasing' throughout this paper.

[2]For any poset $(X, \leq)$, we denote by $\leq_m$ the product order on X^m, i.e., for any $\mathbf{x}, \mathbf{y} \in X^m$, we write $\mathbf{x} \leq_m \mathbf{y}$ if $x_i \leq y_i$ for all $i \in \{1, \ldots, m\}$.

Driven by the existence of interesting functions that are not idempotent yet monotone, averaging functions have been further generalized by replacing the property of idempotence by the property of preservation of the bounds, resulting in the introduction of the notion of an aggregation function [25], [26]. Note that we now need to deal with a *bounded* poset, typically a compact real interval.

Definition 2: Consider a bounded poset $(X, \leq, 0, 1)$ and $n \in \mathbb{N}$. A function $A : X^n \to X$ is called an (n-ary) aggregation function on $(X, \leq, 0, 1)$ if

(i) it satisfies the boundary conditions, i.e., $A(0, \ldots, 0) = 0$ and $A(1, \ldots, 1) = 1$;
(ii) it is monotone.

Although the study of aggregation and averaging functions carries an impressive body of mathematical knowledge, it is quite unfortunate that most studies within the field of aggregation theory are confined to the aggregation of numerical values, i.e., usually a subinterval of the real line (typically $[0, 1]$, $\mathbb{R}^+$ or $\mathbb{R}$ itself) or of values in a linearly ordered (linguistic) scale. Admittedly, the field of application is embarrassingly narrow and the aggregation of many common types of mathematical object, such as rankings, strings and compositional data vectors, does not fit within the current theory of aggregation. In the upcoming section, we propose a natural definition of an aggregation function beyond the current restriction to the poset framework.

III. Aggregation on besets

Given a poset, the notion of an element that is in between two other elements has attracted the attention of many researchers and has been formalized by a mathematical object called betweenness relation [27], [28]. This mathematical object has also been studied given many other types of structures, e.g., metric spaces [29] and road systems [30]. Although many alternative axiomatic definitions have been provided (see, for instance, any of the aforementioned papers), throughout this paper we restrict to the following one, which has already been considered within the context of data aggregation [23].

Definition 3: A ternary relation B on a non-empty set X is called a betweenness relation if it satisfies the following three properties:

(i) *Symmetry in the end points*: for any $x, y, z \in X$, it holds that
$$(x, y, z) \in B \Leftrightarrow (z, y, x) \in B\,.$$
(ii) *Closure*: for any $x, y, z \in X$, it holds that
$$\big((x, y, z) \in B \wedge (x, z, y) \in B\big) \Leftrightarrow y = z\,.$$
(iii) *End-point transitivity*: for any $o, x, y, z \in X$, it holds that
$$\big((o, x, y) \in B \wedge (o, y, z) \in B\big) \Rightarrow (o, x, z) \in B\,.$$

A set X equipped with a betweenness relation B is called a *beset* and denoted by (X, B).

On any set a trivial betweenness relation may be defined.

Example 1: The minimal betweenness relation on a set X is the ternary relation B_0 on X defined as
$$B_0 = \big\{(x, y, z) \in X^3 \mid x = y \vee y = z\big\}\,.$$

Obviously, it is contained in any possible betweenness relation on X.

Nevertheless, one almost surely wants to define a richer betweenness relation than the minimal betweenness relation. Perhaps the two most prominent examples of sets on which a natural betweennes relation could be defined are posets and metric spaces. However, the family of sets on which a natural betweennes relation can be defined is much richer and includes many other structures such as real vector spaces and topological spaces. For further details, we refer to [30].

Example 2: The betweenness relation induced by the order relation $\leq$ of a poset $(X, \leq)$ is the ternary relation $B_{\leq}$ on X defined as

$$B_{\leq} = B_0 \cup \big\{(x, y, z) \in X^3 \mid (x \leq y \leq z) \vee (z \leq y \leq x)\big\} \,.$$

Example 3: The betweenness relation induced by the distance metric d of a metric space (X, d) is the ternary relation B_d on X defined as

$$B_d = \big\{(x, y, z) \in X^3 \mid d(x, z) = d(x, y) + d(y, z)\big\} \,.$$

The product betweenness relation defined on a product set deserves special attention. It will play an important role in the definition of an aggregation function on a beset.

Example 4: Given a betweenness relation B on a set X and $n \in \mathbb{N}$, the product betweenness relation on X^n induced by B is the ternary relation $B^{(n)}$ defined as

$$B^{(n)} = \left\{ (\mathbf{x}, \mathbf{y}, \mathbf{z}) \in (X^n)^3 \;\middle|\; \begin{array}{l} (\forall i \in \{1, \ldots, n\}) \\ ((x_i, y_i, z_i) \in B) \end{array} \right\} \,.$$

Remark 1: Consider a poset $(X, \leq)$. Note that the betweenness relation $B^{(n)}_{\leq}$ on X^n does not need to coincide with the betweenness relation $B_{\leq_n}$. Consider, for instance, X to be the interval $[0, 1]$ and $\leq$ to be the usual order relation on $\mathbb{R}$. Consider $\mathbf{x} = (0, 1)$, $\mathbf{y} = (0.5, 0.5)$ and $\mathbf{z} = (1, 0)$. It follows that $(\mathbf{x}, \mathbf{y}, \mathbf{z}) \in B^{(2)}_{\leq}$, but $(\mathbf{x}, \mathbf{y}, \mathbf{z}) \notin B_{\leq_2}$.

As in the case of posets, one could be interested in the notion of bounds of a beset.

Definition 4: Given a beset (X, B), a non-empty subset S of X is called a set of bounds of (X, B) if, for any $y \in S$ and any $x, z \in X \backslash S$, it holds that $(x, y, z) \notin B$. We thus refer to (X, B, S) as a bounded beset.

Obviously, X is always a set of bounds of any beset (X, B). However, one could note that proper subsets of X may also be a set of bounds of (X, B). For instance, given the betweenness relation $B_{\leq}$ induced by the order relation $\leq$ of a bounded poset $(X, \leq, 0, 1)$, as in Example 2, one could easily verify that $\{0, 1\}$ is also a set of bounds of $(X, B_{\leq})$.

The definition of an aggregation function on a poset could thus be naturally generalized in order to deal with besets.

Definition 5: Consider a bounded beset (X, B, S) and $n \in \mathbb{N}$. A function $A : X^n \to X$ is called an (n-ary) aggregation function on (X, B, S) if

(i) it satisfies the boundary conditions, i.e., $A(o, \ldots, o) = o$, for any $o \in S$;

(ii) it is monotone, i.e., for any $o \in S$ and any $\mathbf{x}, \mathbf{y} \in X^n$, the fact that $\big((o, \ldots, o), \mathbf{x}, \mathbf{y}\big) \in B^{(n)}$ implies that $\big(A(o, \ldots, o), A(\mathbf{x}), A(\mathbf{y})\big) \in B$.

Note that, if $S = X$, then an aggregation function is obviously idempotent. Moreover, if we would allow S to be the empty set, then any function $A : X^n \to X$ would satisfy the definition above (hence, in case $X = [0, 1]$, we would obtain the set of all fusion functions [9]).

The first example of aggregation on besets is that of classical aggregation on posets. Since every poset is a beset, a mandatory requirement for this new definition of an aggregation function is that it needs to coincide with the standard definition of an aggregation function when restricted to posets.

Theorem 1: Consider a bounded poset $(X, \leq, 0, 1)$, the associated bounded beset $(X, B_{\leq}, \{0, 1\})$ and $n \in \mathbb{N}$. A function $A : X^n \to X$ is an (n-ary) aggregation function on $(X, \leq, 0, 1)$ in the sense of Definition 2 if and only if it is an (n-ary) aggregation function on $(X, B_{\leq}, \{0, 1\})$ in the sense of Definition 5.

Proof. Note that the boundary conditions coincide in both definitions. Thus, we only prove that the monotonicity properties are equivalent.

On the one hand, suppose that A is an (n-ary) aggregation function on $(X, \leq, 0, 1)$ in the sense of Definition 2. Consider any $o \in \{0, 1\}$ and any $\mathbf{x}, \mathbf{y} \in X^n$ such that $\big((o, \ldots, o), \mathbf{x}, \mathbf{y}\big) \in B^{(n)}_{\leq}$. If $o = 0$, then it follows that:

$$\big((0, \ldots, 0), \mathbf{x}, \mathbf{y}\big) \in B^{(n)}_{\leq} \,,$$

which implies

$$(0, \ldots, 0) \leq_n \mathbf{x} \leq_n \mathbf{y} \,,$$

and, thus,

$$0 = A(0, \ldots, 0) \leq A(\mathbf{x}) \leq A(\mathbf{y}) \,,$$

and, finally,

$$\big(A(0, \ldots, 0), A(\mathbf{x}), A(\mathbf{y})\big) \in B_{\leq} \,.$$

If $o = 1$, then it follows that:

$$\big((1, \ldots, 1), \mathbf{x}, \mathbf{y}\big) \in B^{(n)}_{\leq} \,,$$

which implies

$$\mathbf{y} \leq_n \mathbf{x} \leq_n (1, \ldots, 1) \,,$$

and, thus,

$$A(\mathbf{y}) \leq A(\mathbf{x}) \leq A(1, \ldots, 1) = 1 \,,$$

and, finally,

$$\big(A(1, \ldots, 1), A(\mathbf{x}), A(\mathbf{y})\big) \in B_{\leq} \,.$$

Thus, A is an (n-ary) aggregation function on $(X, B_{\leq}, \{0, 1\})$ in the sense of Definition 5.

On the other hand, suppose that A is an (n-ary) aggregation function on $(X, B_{\leq}, \{0, 1\})$ in the sense of Definition 5. Consider any $\mathbf{x}, \mathbf{y} \in X^n$ such that $\mathbf{x} \leq_n \mathbf{y}$. Since 0 is the smallest element of X, it holds that $(0, \ldots, 0) \leq_n \mathbf{x}$, and, thus, $\big((0, \ldots, 0), \mathbf{x}, \mathbf{y}\big) \in B^{(n)}_{\leq}$. Since A is an aggregation function on $(X, B_{\leq}, \{0, 1\})$, it follows that $\big(A(0, \ldots, 0), A(\mathbf{x}), A(\mathbf{y})\big) \in B_{\leq}$. Again, since $A(0, \ldots, 0) = 0$ is the smallest element of X, we conclude that $0 \leq A(\mathbf{x}) \leq A(\mathbf{y})$. Thus, A is an (n-ary) aggregation function on $(X, \leq, 0, 1)$ in the sense of Definition 2. □

IV. Examples of aggregation functions on besets

A. Aggregation of compositional data

Think of two liquids of which the ratios of their different compounds are known. If one mixes both liquids in a one-to-one ratio, then the composition of the resulting liquid is known to be given by the componentwise arithmetic mean of the compositions of both original liquids. The result of this mixture should be understood as the result of an aggregation process, however, no intuitive order[3] could be used for defining the monotonicity property. In this subsection, we prove that the aforementioned process is a clear example of aggregation on the (bounded be)set of compositional data vectors.

Vectors of positive real numbers adding up to one are hereinafter referred to as compositional data vectors. Due to their natural interpretation as the proportions of different compounds in a mixture, they are common in many fields of application [31], [32]. Usually, the set of all k-dimensional compositional data vectors is referred to as the k-dimensional simplex and is defined as

$$\mathcal{S}_k = \left\{ \mathbf{x} \in [0,1]^k \;\middle|\; \sum_{j=1}^k \mathbf{x}(j) = 1 \right\}.$$

For the case $k = 3$, the simplex can be naturally represented by an equilateral triangle where the length of each of the medians[4] equals one. Every point inside the equilateral triangle corresponds to a point of the simplex. The coordinates of any point are obtained by the projection of the given point to each of the medians. Figure 1 illustrates the coordinates of the compositional data vector $(0.55, 0.32, 0.13)$.

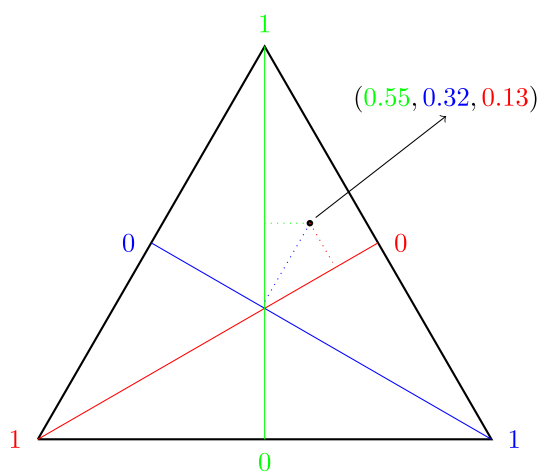

Fig. 1. Graphical representation of the 3-dimensional simplex.

A natural betweenness relation $B_{\mathcal{S}_k}$ on $\mathcal{S}_k$ is defined as follows:

$$B_{\mathcal{S}_k} = \left\{ (\mathbf{x},\mathbf{y},\mathbf{z}) \in (\mathcal{S}_k)^3 \;\middle|\; \begin{array}{l} (\forall j \in \{1,\dots,k\}) \\ (\min(\mathbf{x}(j),\mathbf{z}(j)) \leq \mathbf{y}(j) \leq \max(\mathbf{x}(j),\mathbf{z}(j))) \end{array} \right\}.$$

[3] Admittedly, the Lorenz order [21], which is related to the notion of (stochastic) dominance [22] and commonly used in the field of economics for measuring the concentration of wealth, could be used. Unfortunately, the Lorenz order carries an undesirable behaviour since it assumes symmetry among the different compounds (and it actually is a preorder on the set of compositional data vectors).

[4] A median of a triangle is a line segment joining a vertex to the midpoint of the opposite side.

Figure 2 illustrates the points that are in between two compositional data vectors according to the betweenness relation $B_{\mathcal{S}_k}$. Note that $B_{\mathcal{S}_k}$ is just a reduction of $B_{\leq}^{(k)}$ to triplets of points in the simplex.

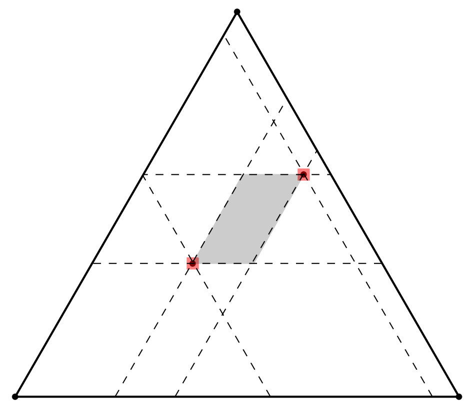

Fig. 2. Illustration of the compositional data vectors (highlighted in grey) that are strictly in between the compositional data vectors that are highlighted in red according to the betweenness relation $B_{\mathcal{S}_3}$.

The aggregation of compositional data vectors appears naturally in many fields of application. Formally, we can aggregate n compositional data vectors $\mathbf{x}_1, \dots, \mathbf{x}_n \in \mathcal{S}_k$, resulting in a new compositional data vector, by using the function $\mathbf{C}_{\mathbf{w}} : (\mathcal{S}_k)^n \to \mathcal{S}_k$ defined by

$$\mathbf{C}_{\mathbf{w}}(\mathbf{x}_1, \dots, \mathbf{x}_n)(j) = \sum_{i=1}^n w_i \mathbf{x}_i(j)\,, \tag{1}$$

for any $j \in \{1, \dots, k\}$, where $\mathbf{w} = (w_1, \dots, w_n)$ is a suitable weighing vector (in the case of the liquids, regarding the mixing ratio associated with each of the different liquids in the mixture).

As an illustrative example, consider the compositional data vectors $\mathbf{x}_1 = (0.55, 0.32, 0.13), \mathbf{x}_2 = (0.03, 0.72, 0.25) \in \mathcal{S}_3$ representing the composition of two liquids in terms of three compounds. If we mix both liquids in a one-to-one ratio, the resulting mixture will have the following composition:

$$\mathbf{C}_{(\frac{1}{2},\frac{1}{2})}(\mathbf{x}_1, \mathbf{x}_2) = (0.29, 0.52, 0.19)\,.$$

In case different quantities are used for each of the liquids, a weighted arithmetic mean, instead of the usual arithmetic mean, needs to be considered. For instance, in case we mix the previous two liquids in such a way that the quantity of $\mathbf{x}_1$ is the triple of the quantity of $\mathbf{x}_2$, the following compositional data vector would be obtained:

$$\mathbf{C}_{(\frac{3}{4},\frac{1}{4})}(\mathbf{x}_1, \mathbf{x}_2) = (0.42, 0.42, 0.16)\,.$$

As intuitive as this sounds, this function is not an aggregation function in the most classical sense. Hereinafter, we prove that it is an aggregation function on a bounded beset.

It is easy to verify that the standard basis of $\mathbb{R}^k$, denoted by $\mathbb{E} = \{\mathbf{e}_j\}_{j=1}^k$, is a set of bounds of the beset $(\mathcal{S}_k, B_{\mathcal{S}_k})$, therefore, $(\mathcal{S}_k, B_{\mathcal{S}_k}, \mathbb{E})$ is a bounded beset. The function defined by Eq. (1), for any weighing vector $\mathbf{w} = (w_1, \dots, w_n)$, is easily proved to be an aggregation function on $(\mathcal{S}_k, B_{\mathcal{S}_k}, \mathbb{E})$:

(i) Consider any $\mathbf{e}_\ell \in \mathbb{E}$. It follows that

$$\mathbf{C}_\mathbf{w}(\mathbf{e}_\ell, \ldots, \mathbf{e}_\ell)(\ell) = \sum_{i=1}^{n} w_i \, 1 = 1 \, .$$

Similarly, for any $j \neq \ell$, it holds that

$$\mathbf{C}_\mathbf{w}(\mathbf{e}_\ell, \ldots, \mathbf{e}_\ell)(j) = \sum_{i=1}^{n} w_i \, 0 = 0 \, .$$

Thus, we conclude that

$$\mathbf{C}_\mathbf{w}(\mathbf{e}_\ell, \ldots, \mathbf{e}_\ell) = \mathbf{e}_\ell \, .$$

(ii) Consider any $\mathbf{e}_\ell \in \mathbb{E}$ and any $(\mathbf{x}_1, \ldots, \mathbf{x}_n)$, $(\mathbf{y}_1, \ldots, \mathbf{y}_n) \in (\mathcal{S}_k)^n$ such that $\big((\mathbf{e}_\ell, \ldots, \mathbf{e}_\ell), (\mathbf{x}_1, \ldots, \mathbf{x}_n), (\mathbf{y}_1, \ldots, \mathbf{y}_n)\big) \in B_{\mathcal{S}_k}^{(n)}$. It follows that $\mathbf{y}_i(\ell) \leq \mathbf{x}_i(\ell) \leq \mathbf{e}_\ell(\ell) = 1$ and $0 = \mathbf{e}_\ell(j) \leq \mathbf{x}_i(j) \leq \mathbf{y}_i(j)$, for any $j \neq \ell$. Thus, it holds that

$$\begin{aligned}\mathbf{C}_\mathbf{w}(\mathbf{y}_1, \ldots, \mathbf{y}_n)(\ell) &\leq \mathbf{C}_\mathbf{w}(\mathbf{x}_1, \ldots, \mathbf{x}_n)(\ell) \\ &\leq \mathbf{C}_\mathbf{w}(\mathbf{e}_\ell, \ldots, \mathbf{e}_\ell)(\ell) \, ,\end{aligned}$$

and that

$$\begin{aligned}\mathbf{C}_\mathbf{w}(\mathbf{e}_\ell, \ldots, \mathbf{e}_\ell)(j) &\leq \mathbf{C}_\mathbf{w}(\mathbf{x}_1, \ldots, \mathbf{x}_n)(j) \\ &\leq \mathbf{C}_\mathbf{w}(\mathbf{y}_1, \ldots, \mathbf{y}_n)(j) \, ,\end{aligned}$$

for any $j \neq \ell$. We conclude that

$$\big(\mathbf{C}_\mathbf{w}(\mathbf{e}_\ell, \ldots, \mathbf{e}_\ell), \mathbf{C}_\mathbf{w}(\mathbf{x}_1, \ldots, \mathbf{x}_n), \mathbf{C}_\mathbf{w}(\mathbf{y}_1, \ldots, \mathbf{y}_n)\big) \in B_{\mathcal{S}_k} \, .$$

B. *Aggregation of binary strings*

Think of several Boolean values to be combined. For instance, consider either the logical 'and' or the logical 'or' operation. Similarly, one could consider the 'mode' operation, which yields a unique aggregate in case an odd number of Boolean values is considered. One could immediately see that the more 0's are changed into 1's, the greater the obtained value for all of these operations. Analogously, the more 1's are changed into 0's, the lower the obtained value. These three operations are indeed monotone w.r.t. the order relation $\leq = \{(0,0),(0,1),(1,1)\}$ according to the standard definition of monotonicity of an aggregation function on a bounded poset. However, one might doubt whether this order-based monotonicity is actually meaningful when the symbols 0 and 1 are substituted by the words 'FALSE' and 'TRUE', or, going to a more extreme case, by an alternative colour-based encoding 'red' and 'green'. This monotonicity property could actually be equivalently defined without assuming any notion of order.

Consider the beset (Σ, B_H, Σ), where $\Sigma = \{a, b\}$ is a binary alphabet and B_H is the betweenness relation on Σ induced by the Hamming distance function[5], i.e.,

$$B_H = \big\{(x, y, z) \in (\Sigma)^3 \mid H(x, z) = H(x, y) + H(y, z)\big\} \, .$$

Note that $B_H = B_0$ in this setting. Interestingly, a function $A : \Sigma^n \to \Sigma$ is monotone if, for any $\mathbf{x}, \mathbf{y} \in \Sigma^n$, it holds that $H(\mathbf{x}, (a, \ldots, a)) = H(\mathbf{x}, \mathbf{y}) + H(\mathbf{y}, (a, \ldots, a))$ implies $A(\mathbf{y}) \in \{A(\mathbf{x}), a\}$ and that $H(\mathbf{x}, (b, \ldots, b)) = H(\mathbf{x}, \mathbf{y}) + H(\mathbf{y}, (b, \ldots, b))$ implies $A(\mathbf{y}) \in \{A(\mathbf{x}), b\}$. This definition coincides with the order-based definition of monotonicity, while it does not assume that $a \leq b$ or, conversely, that $b \leq a$.

This problem can also be considered in an extended form in which, rather than aggregating elements (belonging to the binary alphabet), we aim at aggregating lists of elements. Hereinafter, any list of k elements in Σ (and thus an element in Σ^k) is referred to as a string of length k. Prominent operations on strings are the bitwise extensions of the aforementioned logical 'and', logical 'or' and 'mode' (the latter one assuming k is an odd number).

Again, without assuming any order relation, neither on Σ nor on Σ^k, we can still define the monotonicity of a function on the bounded beset $(\Sigma^k, B_H, \Sigma^k)$, where B_H is now the following betweenness relation on Σ^k:

$$B_H = \big\{(\mathbf{x}, \mathbf{y}, \mathbf{z}) \in (\Sigma^k)^3 \mid H(\mathbf{x}, \mathbf{z}) = H(\mathbf{x}, \mathbf{y}) + H(\mathbf{y}, \mathbf{z})\big\} \, .$$

Note that $B_H \neq B_0$ in case $k > 1$. Figure 3 illustrates an example of the strings that are in between 'aab' and 'bba' according to the betweenness relation B_H for the binary alphabet $\Sigma = \{a, b\}$.

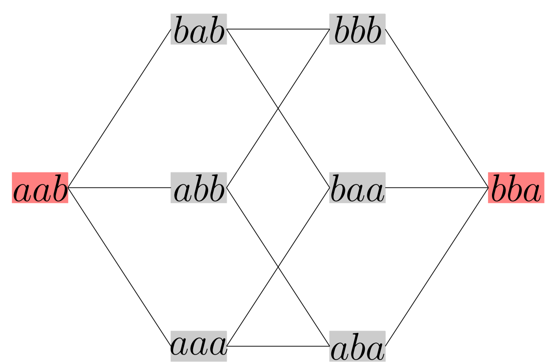

Fig. 3. Illustration of the strings (highlighted in grey) that are strictly in between the strings aab and bba (highlighted in red) according to the betweenness relation B_H for the binary alphabet $\Sigma = \{a, b\}$.

Any among the three bitwise extensions of the logical 'and', logical 'or'[6] and 'mode' is easily proved to be an aggregation function on $(\Sigma^k, B_H, \Sigma^k)$ due to their idempotence and bitwise nature.

V. Prominent monotonicity failures

In this section, we analyse two common examples of aggregation processes that do not fulfill the property of monotonicity as defined in Section III: the aggregation of m-ary strings and the aggregation of rankings. Note that, in both settings, it is common that the aggregation of n objects might not be uniquely defined. In order to ease the remainder of the paper, we will restrict to illustrative examples in which the result of the aggregation is unique.

[5] Given a finite alphabet Σ and a natural number n, the Hamming distance function $H : \Sigma^n \times \Sigma^n \to \mathbb{R}$ is defined as follows: $H(\mathbf{x}, \mathbf{y})$ equals the number of positions at which the lists $\mathbf{x}$ and $\mathbf{y}$ differ [33].

[6] Assuming again $a = 0$ and $b = 1$.

A. Aggregation of m-ary strings

As discussed in Subsection IV-B, the bitwise extensions of the logical 'and', the logical 'or' and the 'mode' are all examples of aggregation functions for binary strings in the sense of Definition 5. However, when m-ary strings (lists of elements in an alphabet of cardinality $m > 2$) are considered, the monotonicity property might no longer be satisfied as we see in the following example. Consider the alphabet $\Sigma = \{a, b, c\}$, then the positionwise[7] extension of the 'mode' is no longer an aggregation function on $(\Sigma^k, B_H, \Sigma^k)$. For instance, consider $k = 1$ and $n = 7$, then one could see that $((a, a, a, a, b, b, b), (a, a, c, c, b, b, b), (c, c, c, c, c, c, c)) \in B_H^{(7)}$, whereas $(a, b, c) \notin B_H$.

B. Aggregation of rankings

The aggregation of rankings is a popular topic that has been addressed in many scientific disciplines including medicine [34], consumer preference analysis [35], computer science [36] and, mainly, social choice theory [37], [38].

Formally, a ranking (without ties) is a strict total order relation $\succ$ on a set $\mathscr{C} = \{a_1, \dots, a_k\}$ of k elements, i.e., the asymmetric part of a total order relation $\succeq$ on $\mathscr{C}$. The set of all rankings on $\mathscr{C}$ is denoted by $\mathcal{L}(\mathscr{C})$. The most common notion of distance on rankings is measured by means of the Kendall distance function K between rankings [39]. This distance function assigns to each couple of rankings the number of pairwise disagreements between them, i.e., for any two rankings $\succ_1$ and $\succ_2$, the Kendall distance is defined as

$$K(\succ_1, \succ_2) = \#\{(a_{i_1}, a_{i_2}) \in \mathscr{C}^2_{\neq} \mid a_{i_1} \succ_1 a_{i_2} \wedge a_{i_2} \succ_2 a_{i_1}\}\,.$$

The Kendall distance function induces a natural betweenness relation B_K on $\mathcal{L}(\mathscr{C})$, which is defined as follows:

$$B_K = \{(\succ_1, \succ_2, \succ_3) \in \mathcal{L}(\mathscr{C})^3 \mid K(\succ_1, \succ_3) = K(\succ_1, \succ_2) + K(\succ_2, \succ_3)\}\,.$$

Figure 4 illustrates the rankings that are in between the rankings $a \succ b \succ c \succ d$ and $d \succ b \succ a \succ c$ according to the betweenness relation B_K for the set $\mathscr{C} = \{a, b, c, d\}$.

Probably the most prominent method for the aggregation of rankings is that of Kemeny [40], which selects as the aggregate of a list of rankings $\mathscr{R} = (\succ_i)_{i=1}^n \in \mathcal{L}(\mathscr{C})^n$, the ranking $\succ \in \mathcal{L}(\mathscr{C})$ that minimizes the sum of Kendall distances to $\mathscr{R}$, i.e.,

$$\underset{\succ \in \mathcal{L}(\mathscr{C})}{\arg\min} \sum_{i=1}^{n} K(\succ_i, \succ)\,.$$

Unfortunately, the method of Kemeny is not an aggregation function on $(\mathcal{L}(\mathscr{C}), B_K, \mathcal{L}(\mathscr{C}))$. For instance, consider $\mathscr{C} = \{a, b, c\}$ and the lists of $n = 11$ rankings $\mathscr{R}$, $\mathscr{R}'$ and $\mathscr{R}''$ shown in Table I. The result of aggregating $\mathscr{R}$, $\mathscr{R}'$ and $\mathscr{R}''$ by means of the method of Kemeny is $a \succ b \succ c$, $b \succ c \succ a$ and $c \succ a \succ b$, respectively. One could verify that $(\mathscr{R}, \mathscr{R}', \mathscr{R}'') \in B_K^{(11)}$, whereas $(a \succ b \succ c, b \succ c \succ a, c \succ a \succ b) \notin B_K$.

[7]When talking about m-ary strings we rather use the term 'positionwise' instead of 'bitwise'.

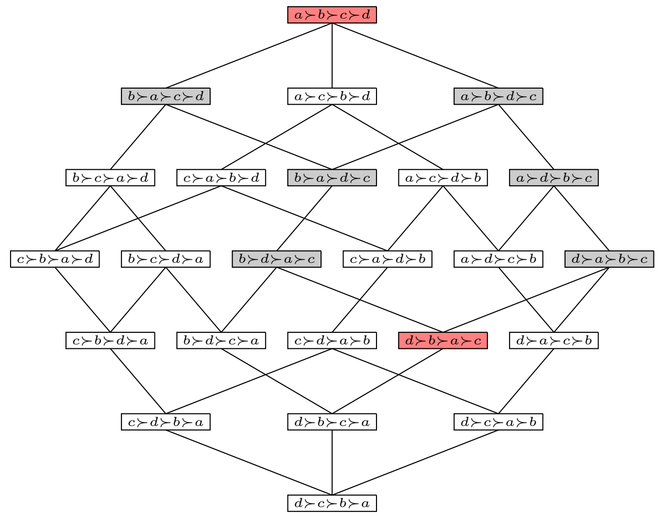

Fig. 4. Illustration of the rankings (highlighted in grey) that are strictly in between the rankings $a \succ b \succ c \succ d$ and $d \succ b \succ a \succ c$ (highlighted in red) according to the betweenness relation B_K.

$\mathscr{R}$	$\mathscr{R}'$	$\mathscr{R}''$
$a \succ b \succ c$	$c \succ a \succ b$	$c \succ a \succ b$
$a \succ b \succ c$	$c \succ a \succ b$	$c \succ a \succ b$
$a \succ b \succ c$	$a \succ b \succ c$	$c \succ a \succ b$
$a \succ b \succ c$	$a \succ b \succ c$	$c \succ a \succ b$
$b \succ a \succ c$	$b \succ c \succ a$	$c \succ a \succ b$
$b \succ a \succ c$	$b \succ c \succ a$	$c \succ a \succ b$
$b \succ a \succ c$	$b \succ c \succ a$	$c \succ a \succ b$
$b \succ a \succ c$	$b \succ c \succ a$	$c \succ a \succ b$
$c \succ a \succ b$	$c \succ a \succ b$	$c \succ a \succ b$
$c \succ a \succ b$	$c \succ a \succ b$	$c \succ a \succ b$
$b \succ c \succ a$	$b \succ c \succ a$	$c \succ a \succ b$

TABLE I
LISTS $\mathscr{R}$, $\mathscr{R}'$ AND $\mathscr{R}''$ OF RANKINGS ON $\mathscr{C} = \{a, b, c\}$.

VI. DISCUSSION

In this paper, we have presented a generalization of the definition of an aggregation function that goes beyond the current restriction to posets. Unfortunately, we have found two prominent examples of aggregation processes that do not fit within this newly-introduced definition. In particular, they fail to fulfill the property of monotonicity. A weaker version of the property of monotonicity embracing these two examples could be thought of. Instead of requiring that if the elements to be aggregated 'move towards' any bound, then the result of the aggregation also 'moves towards' this very same bound, we could soften this requirement as follows: if the result of the aggregation is a bound and the elements to be aggregated 'move towards' this bound, then the result of the aggregation should not change.

Definition 6: Consider a bounded beset (X, B, S) and $n \in \mathbb{N}$. A function $A : X^n \to X$ is called quasimonotone if for any $o \in S$ and any $\mathbf{x}, \mathbf{y} \in X^n$, the facts that $\big((o, \dots, o), \mathbf{x}, \mathbf{y}\big) \in B^{(n)}$ and $A(o, \dots, o) = A(\mathbf{y})$ jointly imply that $A(o, \dots, o) = A(\mathbf{x})$.

It is straightforward to see that the positionwise mode for the aggregation of strings satisfies quasimonotonicity, even when dealing with alphabets of cardinality greater than two. Additionally, it is easily verified that the method of Kemeny for the aggregation of rankings also satisfies quasimonotonicity. The proof could be sketched as follows. Consider a list of rankings $(\succ_1, \ldots, \succ_n) \in \mathcal{L}(\mathscr{C})^n$ to be aggregated and suppose $\succ^*$ is the unique Kemeny ranking, i.e.,

$$\succ^* = \arg\min_{\succ \in \mathcal{L}(\mathscr{C})} \sum_{i=1}^{n} K(\succ_i, \succ) \, .$$

Consider $(\succ_1', \ldots, \succ_n') \in \mathcal{L}(\mathscr{C})^n$ such that[8]

$$\sum_{i=1}^{n} K(\succ_i', \succ^*) = \sum_{i=1}^{n} K(\succ_i, \succ^*) - 1$$

and

$$\big((\succ^*, \ldots, \succ^*), (\succ_1', \ldots, \succ_n'), (\succ_1, \ldots, \succ_n)\big) \in B_K^{(n)} \, .$$

Note that, for any $\succ \neq \succ^*$, it holds that

$$\sum_{i=1}^{n} K(\succ_i', \succ) \in \left\{ \sum_{i=1}^{n} K(\succ_i, \succ) - 1, \sum_{i=1}^{n} K(\succ_i, \succ) + 1 \right\} \, .$$

Thus, $\succ^*$ remains being the unique Kemeny ranking for $(\succ_1', \ldots, \succ_n')$. Proceeding in an iterative manner, the property of quasimonotonicity is easily verified to hold.

Although the property of quasimonotonicity seems to be quite intuitive when dealing with the aggregation of m-ary strings or the aggregation of rankings, it may turn out to be too weak when moving to other settings, e.g., the classical aggregation of real numbers. More precisely, given the bounded beset $([a, b], \leq, \{a, b\})$, any function $f : [a, b]^n \to [a, b]$ verifying that (1) $f(\mathbf{x}) = a$ if and only if $\mathbf{x} = (a, \ldots, a)$ and (2) $f(\mathbf{x}) = b$ if and only if $\mathbf{x} = (b, \ldots, b)$, is quasimonotone.

We end by concluding that, similarly to non-symmetric functions (such as weighted arithmetic means) and non-idempotent functions (such as copulas), non-monotone functions should also be accommodated within the field of aggregation theory. In this direction, penalty-based functions have been deeply studied for the aggregation of real numbers [41], [42] and, more generally, for the aggregation on besets [23].

References

[1] M. de Carvalho, "Mean, what do you mean?," *The American Statistician*, vol. 70, no. 3, pp. 270–274, 2016.

[2] A. L. Cauchy, *Cours d'analyse de l'Ecole Royale Polytechnique*, vol. 1 of *Analyse algébraique*. Paris: Debure, 1821.

[3] M. Grabisch, J.-L. Marichal, R. Mesiar, and E. Pap, "Aggregation functions: Means," *Information Sciences*, vol. 181, pp. 1–22, 2011.

[4] D. Dubois and H. Prade, "On the use of aggregation operations in information fusion processes," *Fuzzy Sets and Systems*, vol. 142, pp. 143–161, 2004.

[5] M. Komorníková and R. Mesiar, "Aggregation functions on bounded partially ordered sets and their classification," *Fuzzy Sets and Systems*, vol. 175, pp. 48–56, 2011.

[6] T. Wilkin and G. Beliakov, "Weakly monotonic averaging functions," *International Journal of Intelligent Systems*, vol. 30, no. 2, pp. 144–169, 2015.

[7] G. Beliakov and J. Špirková, "Weak monotonicity of Lehmer and Gini means," *Fuzzy Sets and Systems*, vol. 299, pp. 26–40, 2016.

[8] Note that such list of rankings is easily obtained by reversing exactly one couple (a_{i_1}, a_{i_2}) of consecutive elements in exactly one of the rankings $\succ_i$ (with $\succ_i \neq \succ^*$) such that it now holds that $a_{i_2} \succ a_{i_1}$, $a_{i_1} \succ' a_{i_2}$ and $a_{i_1} \succ^* a_{i_2}$. Obviously, this is no longer possible when $(\succ_1, \ldots, \succ_n) = (\succ^*, \ldots, \succ^*)$.

[8] G. Beliakov, T. Calvo, and T. Wilkin, "On the weak monotonicity of Gini means and other mixture functions," *Information Sciences*, vol. 300, pp. 70–84, 2015.

[9] H. Bustince, J. Fernández, A. Kolesárová, and R. Mesiar, "Directional monotonicity of fusion functions," *European Journal of Operational Research*, vol. 244, pp. 300–308, 2015.

[10] M. Gagolewski, "Penalty-based aggregation of multidimensional data," *Fuzzy Sets and Systems*, vol. 325, pp. 4–20, 2017.

[11] M. Gagolewski, R. Pérez-Fernández, and B. De Baets, "An inherent difficulty in the aggregation of multidimensional data," *IEEE Transactions on Fuzzy Systems, in press, DOI: 10.1109/TFUZZ.2019.2908135*.

[12] R. Pérez-Fernández, B. De Baets, and M. Gagolewski, "A taxonomy of monotonicity properties for the aggregation of multidimensional data," *Information Fusion*, vol. 52, pp. 322–334, 2019.

[13] K. J. Arrow, *Social Choice and Individual Values*. New Haven: Yale University Press, second ed., 1963.

[14] J. H. Smith, "Aggregation of preferences with variable electorate," *Econometrica*, vol. 41, pp. 1027–1041, 1973.

[15] P. C. Fishburn, "Monotonicity paradoxes in the theory of elections," *Discrete Applied Mathematics*, vol. 4, no. 2, pp. 119–134, 1982.

[16] J. P. Barthelemy and B. Monjardet, "The median procedure in cluster analysis and social choice theory," *Mathematical Social Sciences*, vol. 1, pp. 235–267, 1981.

[17] B. Monjardet, "An axiomatic theory of tournament aggregation," *Mathematics of Operations Research*, vol. 3, no. 4, pp. 334–351, 1978.

[18] Y. Wakabayashi, "The complexity of computing medians of relations," *Resenhas*, vol. 3, no. 3, pp. 323–349, 1998.

[19] T. Kohonen, "Median strings," *Pattern Recognition Letters*, vol. 3, pp. 309–313, 1985.

[20] J. K. Lanctot, M. Li, B. Ma, S. Wang, and L. Zhang, "Distinguishing string selection problems," *Information and Computation*, vol. 185, pp. 41–55, 2003.

[21] M. O. Lorenz, "Methods of measuring the concentration of wealth," *Publications of the American Statistical Association*, vol. 9, no. 70, pp. 209–219, 1905.

[22] H. Levy, *Stochastic Dominance: Investment Decision Making under Uncertainty*. Berlin: Springer, third ed., 2016.

[23] R. Pérez-Fernández and B. De Baets, "On the role of monometrics in penalty-based data aggregation," *IEEE Transactions on Fuzzy Systems*, vol. 27, no. 7, pp. 1456–1468, 2019.

[24] G. Beliakov, H. Bustince, and T. Calvo, *A Practical Guide to Averaging Functions*. Studies in Fuzziness and Soft Computing, Berlin: Springer-Verlag, Berlin, Heidelberg, 2016.

[25] G. Beliakov, A. Pradera, and T. Calvo, *Aggregation Functions: A Guide for Practitioners*, vol. 221 of *Studies in Fuzziness and Soft Computing*. Berlin, Heidelberg: Springer, 2007.

[26] M. Grabisch, J.-L. Marichal, R. Mesiar, and E. Pap, *Aggregation Functions*. Cambridge: Cambridge University Press, 2009.

[27] P. C. Fishburn, "Betweenness, orders and interval graphs," *Journal of Pure and Applied Algebra*, vol. 1, no. 2, pp. 159–178, 1971.

[28] E. V. Huntington and J. R. Kline, "Sets of independent postulates for betweenness," *Transactions of the American Mathematical Society*, vol. 18, no. 3, pp. 301–325, 1917.

[29] E. Pitcher and M. F. Smiley, "Transitivities of betweenness," *Transactions of the American Mathematical Society*, vol. 52, no. 1, pp. 95–114, 1942.

[30] P. Bankston, "Road systems and betweenness," *Bulletin of Mathematical Sciences*, vol. 3, no. 3, pp. 389–408, 2013.

[31] J. Aitchison, "The statistical analysis of compositional data (with discussion)," *Journal of the Royal Statistical Society*, vol. B44, pp. 139–177, 1982.

[32] J. Aitchison, *The Statistical Analysis of Compositional Data*. Blackburn Press, 2003.

[33] R. W. Hamming, "Error detecting and error correcting codes," *The Bell System Technical Journal*, vol. 29, no. 2, pp. 147–160, 1950.

[34] S. Lin, "Rank aggregation methods," *Wiley Interdisciplinary Reviews: Computational Statistics*, vol. 2, no. 5, pp. 555–570, 2010.

[35] P. Varela and A. Gastón, "Sensory profiling, the blurred line between sensory and consumer science. a review of novel methods for product characterization," *Food Research International*, vol. 48, pp. 893–908, 2012.

[36] C. Dwork, R. Kumar, M. Naor, and D. Sivakumar, "Rank aggregation methods for the web," in *Proceedings of the 10th International Conference on World Wide Web*, pp. 613–622, 2001.

[37] K. J. Arrow, A. K. Sen, and K. Suzumura, *Handbook of Social Choice and Welfare*, vol. 1. Amsterdam: North-Holland, 2002.

[38] D. Black, *The Theory of Committees and Elections*. Cambridge: Cambridge University Press, 1958.
[39] M. G. Kendall, "A new measure of rank correlation," *Biometrika*, vol. 30, pp. 81–93, 1938.
[40] J. G. Kemeny, "Mathematics without numbers," *Daedalus*, vol. 88, no. 4, pp. 577–591, 1959.
[41] T. Calvo and G. Beliakov, "Aggregation functions based on penalties," *Fuzzy Sets and Systems*, vol. 161, pp. 1420–1436, 2010.
[42] H. Bustince, G. Beliakov, G. P. Dimuro, B. Bedregal, and R. Mesiar, "On the definition of penalty functions in data aggregation," *Fuzzy Sets and Systems*, vol. 323, pp. 1–18, 2017.

Raúl Pérez-Fernández holds a B.Sc. in Mathematics (University of Oviedo, Spain, 2012), an M.Sc. in Mathematical Modelling, Statistics and Computing (University of Oviedo, Spain, 2013), and a joint Ph.D. in Mathematics and Statistics (University of Oviedo, Spain, 2017) and Applied Biological Sciences (Ghent University, Belgium, 2017). Currently, he is an FWO postdoctoral fellow at the research unit KERMIT and an Assistant Professor at the University of Oviedo, Spain. His main research interests include aggregation theory and social choice theory.

Bernard De Baets holds an M.Sc. in Maths (1988), a Postgraduate degree in Knowledge Technology (1991) and a Ph.D. in Maths (1995), all summa cum laude from Ghent University (Belgium). He is a Senior Full Professor in Applied Maths (1999) at Ghent University, where he is leading KERMIT, the research unit Knowledge-Based Systems. He is a Government of Canada Award holder (1988), an Honorary Professor of Budapest Tech (2006), an IFSA Fellow (2011), an Honorary Doctor of the University of Turku (2017), and a Profesor Invitado of the Universidad Central "Marta Abreu" de las Villas (2017). He received the 2019 EUSFLAT Scientific Excellence Award. His publications comprise more than 500 papers in international journals and about 60 book chapters. He serves on the Editorial Boards of various international journals, in particular as co-editor-in-chief of Fuzzy Sets and Systems.